PATENT
DOCKET NO. BM45330

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

BOX PCT/DO-EO
Commissioner for Patents
Washington, D.C. 20231

"Express Mail" mailing label number **EK 295 392 146 US**

I hereby certify that this paper is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" under 37 CFR § 1.10 on the date indicated above and addressed to: Commissioner of Patents, Washington, DC 20231, on February 14, 2001.

Colleen Hanagan

NATIONAL STAGE APPLICATION TRANSMITTAL LETTER
APPLICATION FILING UNDER 35 U.S.C. § 371

Transmitted herewith for filing is the patent application of:

Inventor(s)/Applicant(s):	**Thonnard, Joelle**
International Application No.:	**PCT/EP99/05989**
International Published Appln. No.:	**WO 00/11182**
International Filing Date:	**13 August 1999**
Priority Filing Date:	**18 August 1998**
Thirty Month Date:	**18 February 2001**
Title:	**"BASB024 OUTER MEMBRANE PROTEIN OF NEISSERIA MENINGITIDIS"**

1. **THIS NEW APPLICATION IS A NATIONAL STAGE APPLICATION UNDER PCT, CHAPTER II WITH A REQUEST FOR EXAMINATION WITHOUT DELAY TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US).**

☒ This is a **FIRST** submission of items concerning a filing under 35 U.S.C. § 371;

☐ This is a **SECOND** or subsequent submission of items concerning a filing under 35 U.S.C. § 371.

2. ☒ This is an express request to begin national examination procedures (35 U.S.C. 371(f)) at any time rather than delay examination until the expiration of the applicable time limit 35 U.S.C. 371(b) and PCT Articles 22 and 39(1).

3. A proper Demand for International Preliminary Examination was made by the 19TH month from the earliest claimed priority date.

4. Enclosed items are required for filing under 37 CFR § 1.53(b) and § 1.494(b) or § 1.495(b):

☒ One copy of International Publication **WO 00/11182**

(a) ☒ is transmitted herewith **(required only if not transmitted by the International Bureau)**

(b) ☒ has been transmitted by the International Bureau

(c) ☐ is not required, as the application was filed in the United States Receiving Office (RO/US)

Fees

☒ The basic national fee set forth in 37 CFR § 1.482 - International Preliminary Examination Fee not paid to USPTO but International Search Report prepared by the EPO or JPO - **$860.00**

☒ Claims in Excess of 20 (20 @ $18.00)

☒ Independent Claims in Excess of 3 (4 @ $80.00)

5. Further enclosed are:

☒ One copy of International Preliminary Examination Report.

☒ One copy of International Search Report - (references not available).

☐ One copy of Written Opinion.

☐ One copy of PCT Request as filed.

☐ One copy of Chapter II Demand as filed.

6. ☐ A translation of the International Application into English (35 U.S.C. § 371(c)(2))

7. ☐ Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. § 371(c)(3))

8. ☐ A translation of the amendments to the claims under PCT Article 19 (35 U.S.C. § 371(c)(3))

9. Still additional papers enclosed:

☐ Assignment with Assignment Recordation Form Cover Sheet

☐ Verification Statement Claiming Small Entity Status

☐ Declaration or oath is enclosed executed by the inventor

☐ An Information Disclosure Statement under 37 CFR § 1.97 and § 1.98

☒ Return Acknowledgment Postcard

10. Nucleotide and/or Amino Acid Sequence Submission (if applicable, all necessary)

A. Enclosed are:

(a) ☒ Computer Readable Copy of the Sequence Listing

(b) ☒ Paper Copy (identical to Computer Readable Copy) of the Sequence Listing

B. ☐ Enclosed is a paper copy of the Sequence Listing. This paper copy and a Computer Readable Form thereof are identical with the Computer Readable Form in another application of the Applicant which is fully identified as follows:

U.S. Application No.: @@

Filed: @@

Attorney Docket No.: @@

which is believed to comply with the rules set forth in 37 CFR § 1.821 et. seq. Applicants requests pursuant to 37 CFR § 1.821(e) that this Computer Readable Form be used in the present application. **Please TRANSFER the sequence listing from the parent to this application.**

C. ☒ Statement under 37 CFR § 1.821(f): **The information recorded in computer readable form is identical to the written Sequence Listing.**

D. ☒ Statement under 37 CFR § 1.821(g) (required when Sequence Listing not submitted at the time of filing under 35 U.S.C. §111(a)) or 37 CFR §1.821(f) (required when Sequence Listing not submitted at the time of filing under the Patent Cooperation Treaty): **The submission of the Sequence Listing includes no new matter.**

E. ☒ Amendment: Please enter the Sequence Listing into the application.

11. **Preliminary Amendment**

Prior to calculation of fees, kindly enter:

☒ Preliminary Amendment submitted herewith

☐ do not enter Preliminary Amendment

12. Fee payment being made at this time is enclosed:

*	Basic filing fee ($860.00)	860.00
*	Claims in Excess of 20 (20 @ $18.00)	360.00
	Independent Claims in Excess of 3 (4 @ $80.00)	240.00
*	Total Fees enclosed:	$1460.00

13. The Commissioner is hereby authorized to charge any fees which may be required, or credit any overpayment to Account No. 50-0258. This letter is filed in duplicate for accounting purposes.

Respectfully submitted,

Date: February 13, 2001

Arthur E. Jackson
Registration No. 34,354
Allen Bloom
Registration No. 29,135
Attorney for Applicant

DECHERT
4000 Bell Atlantic Tower
1717 Arch Street
Philadelphia, PA 19103-2793
Fax: (609) 620-3259
Attn: Arthur E. Jackson, Esq.
(609) 620-3254

105886.1.07

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant: Thonnard
Serial No.: Unknown
Filed: Herewith

Group Art Unit No.: Unknown
Examiner: Unknown

For: BASB024 Outer Membrane Protein of Neisseria Meningitidis

PRELIMINARY AMENDMENT

Sir:

Please enter this Preliminary Amendment into the record of the above-identified patent application before the calculation of fees.

In the Claims:

Please delete the claims of the application as filed in the PCT and substitute therefor:

29. An isolated polypeptide comprising a member selected from the group consisting of

(a) an amino acid sequence which has at least 90% identity to SEQ ID NOs:4 or 6;

(b) an immunogenic fragment of the amino acid sequence of (a), wherein the immunogenic fragment is at least 90% identical to an aligned contiguous segment of SEQ ID NOs:4 or 6; and

(c) an immunogenic fragment of the amino acid sequence of (a) that matches an aligned contiguous segment of SEQ ID NOs:4 or 6 with no more than five single amino acid substitutions, deletions or additions;

wherein the isolated polypeptide, when administered to a subject in a suitable composition which can include an adjuvant, or a suitable carrier coupled to the polypeptide, induces an immune response that recognizes a polypeptide having the sequence of SEQ ID NOs:4 or 6.

30. The isolated polypeptide of claim 29, wherein the immunogenic fragment of (b) comprises at least 15 amino acids.

31. The isolated polypeptide of claim 30, wherein the immunogenic fragment matches an aligned contiguous segment of SEQ ID NOs:4 or 6 with no more than a single amino acid substitution, deletion or addition.

32. The isolated polypeptide of claim 29, wherein the immunogenic fragment of (b) comprises at least 20 amino acids.

33. The isolated polypeptide of Claim 29 wherein the amino acid sequence of (a) has at least 95% identity to SEQ ID NOs:4 or 6.

34. The isolated polypeptide of Claim 33 wherein the isolated polypeptide comprises the amino acid sequence of SEQ ID NOs:4 or 6.

35. The isolated polypeptide of claim 33 wherein the isolated polypeptide consists of the amino acid sequence of SEQ ID NOs:4 or 6.

36 An isolated polypeptide consisting of the amino acid sequence of SEQ ID NO:2.

37. A fusion protein comprising the isolated polypeptide of Claim 29.

38. A fusion protein comprising the isolated polypeptide of 36.

39. The isolated polypeptide of Claim 29 wherein the polypeptide is the immunogenic fragment having no more than two single amino acid substitutions, deletions or additions relative to the aligned sequence.

40. The isolated polypeptide of Claim 29 wherein the polypeptide is the immunogenic fragment having no more than one single amino acid substitution, deletion or addition relative to the aligned sequence.

41. The isolated polypeptide of Claim 29 wherein the polypeptide is the immunogenic fragment which matches the aligned sequence.

42. An isolated polypeptide encoded by an isolated first polynucleotide wherein the isolated first polynucleotide hybridizes under stringent conditions to a second polynucleotide which encodes the polypeptide of SEQ ID NOs:4 or 6; wherein stringent conditions comprise overnight incubation at 42° C in a solution comprising: 50% formamide, 5×SSC (150 mM

NaCl, 15 mM trisodium citrate), 50 mM sodium phosphate (pH7.6), 5× Denhardt's solution, 10% dextran sulfate, and 20 micrograms/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1× SSC at about 65° C; wherein the isolated polypeptide, when administered to a subject in a suitable composition which can include an adjuvant, or a suitable carrier coupled to the polypeptide, induces an immune response that recognizes a polypeptide having the sequence of SEQ ID NOs:4 or 6.

43. An isolated polynucleotide encoding a polypeptide of Claim 29 or the full complement to the isolated polynucleotide.

44. An isolated polynucleotide encoding a polypeptide of Claim 29, wherein the isolated polynucleotide encodes the polypeptide comprising SEQ ID NOs:4 or 6.

45. An isolated polynucleotide comprising the polynucleotide of SEQ ID NOs:3 or 5.

46. An isolated polynucleotide comprising the polynucleotide of SEQ ID NO:1.

47. An isolated polynucleotide segment comprising a polynucleotide sequence or the full complement of the entire length of the polynucleotide sequence, wherein the polynucleotide sequence hybridizes to the full complement of SEQ ID NOs:3 or 5 minus the full complement of any terminal stop codon, wherein the hybridization conditions include incubation at 42°C in a solution comprising: 50% formamide, 5x SSC (150mM NaCl, 15mM trisodium citrate), 50 mM sodium phosphate (pH7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 micrograms/ml denatured, sheared salmon sperm DNA, followed by washing in 0.1x SSC at 65°C; and, wherein the polynucleotide sequence is identical to SEQ ID NOs:3 or 5 minus any terminal stop codon, except that, over the entire length corresponding to SEQ ID NO:3 or 5 minus any terminal stop codon, $\mathbf{n_n}$ nucleotides are substituted, inserted or deleted, wherein $\mathbf{n_n}$ satisfies the following expression

$$\mathbf{n_n \leq x_n - (x_n \bullet y)}$$

wherein $\mathbf{x_n}$ is the total number of nucleotides in SEQ ID NOs:3 or 5 minus any terminal stop codon, $\mathbf{y}$ is at least 0.95, and wherein any non-integer product of $\mathbf{x_n}$ and $\mathbf{y}$ is rounded down to the nearest integer before subtracting the product from $\mathbf{x_n}$; and wherein the polynucleotide sequence detects *Neisseria meningitidis*.

48. An expression vector comprising the isolated polynucleotide of Claim 43.

49. A host cell transformed with the expression vector of Claim 48.

50. A process of producing an isolated polypeptide comprising (a) culturing the host cell of Claim 49 under conditions sufficient for the production of the encoded polypeptide and (b) recovering the polypeptide.

51. A nucleic acid vaccine comprising the isolated polynucleotide of Claim 43 and a pharmaceutically acceptable carrier.

52. An isolated polynucleotide segment comprising a polynucleotide sequence or the full complement of the entire length of the polynucleotide sequence, wherein the polynucleotide sequence is identical to SEQ ID NOs:3 or 5 minus any terminal stop codon, except that, over the entire length corresponding to SEQ ID NOs:3 or 5 minus any terminal stop codon, $\mathbf{n_n}$ nucleotides are substituted, inserted or deleted, wherein $\mathbf{n_n}$ satisfies the following expression

$$\mathbf{n_n \leq x_n - (x_n \bullet y)}$$

wherein $\mathbf{x_n}$ is the total number of nucleotides in SEQ ID NOs:3 or 5 minus any terminal stop codon, **y** is at least 0.95, and wherein any non-integer product of $\mathbf{x_n}$ and **y** is rounded down to the nearest integer before subtracting the product from $\mathbf{x_n}$; and wherein the polynucleotide sequence detects *Neisseria meningitidis*.

53. The isolated polynucleotide of Claim 52 where **y** is at least 0.97.

54. The isolated polynucleotide of Claim 52, where **y** is at least 0.99.

55. An expression vector comprising the isolated polynucleotide of Claim 52 which codes for a polypeptide that, when administered to a mammal which can include an adjuvant, or a suitable carrier coupled to the polypeptide, induces an immune response that recognizes a polypeptide having the sequence of SEQ ID NOs:4 or 6.

56. A host cell transformed with the isolated polynucleotide or an expression vector comprising the isolated polynucleotide of Claim 52.

57. A process of producing an isolated polypeptide comprising (a) culturing the host cell of Claim 56 under conditions sufficient for the production of the encoded polypeptide and (b) recovering the polypeptide.

58. A vaccine comprising the polypeptide of Claim 29 and a pharmaceutically acceptable carrier.

59. A vaccine comprising the polypeptide of Claim 36 and a pharmaceutically acceptable carrier.

60. The vaccine of Claim 58, wherein the composition comprises at least one other *Neisseria meningitidis* antigen.

61. An antibody immunospecific for the polypeptide or immunogenic fragment of Claim 29.

62. An antibody immunospecific for the polypeptide of Claim 36.

63. A method for inducing an immune response in a mammal comprising administration of the polypeptide of Claim 29.

64. A method of diagnosing a *Neisseria meningitidis* infection, comprising identifying a polypeptide of Claim 29, or an antibody that is immunospecific for the polypeptide, present within a biological sample from an animal suspected of having such an infection.

65. A method for inducing an immune response in a mammal comprising administration of the isolated polynucleotide of Claim 43.

66. A therapeutic composition useful in treating humans with *Neisseria meningitidis* comprising at least one antibody directed against the polypeptide of claim 29 and a suitable pharmaceutical carrier.

67. A therapeutic composition useful in treating humans with *Neisseria meningitidis* comprising at least one antibody directed against the polypeptide of claim 36 and a suitable pharmaceutical carrier.

68. A process for expressing the polynucleotide of Claim 43 comprising transforming a host cell with the expression vector comprising the polynucleotide and culturing the host cell under conditions sufficient for expression of the polynucleotide.

REMARKS

Applicant respectfully requests that this Preliminary Amendment be entered in this case before the calculation of fees and before examination of the subject application.

Claims

Claims 1-28 have been canceled without prejudice or disclaimer of the subject matter therein. Applicant reserves the right to prosecute, in one or more patent applications, the canceled claims, the claims to non-elected inventions, the claims as originally filed, and any other claims supported by the specification.

New claims 29-68 have been introduced. No new matter is added.

Support

Support for the new claims is either obvious, or is as described in the text below. Particularly, support for the recitation of "five single amino acid substitutions, deletions or additions" can be found, for example, at page 6, lines 17-18. Support for compositions of the isolated polypeptide which include an adjuvant recited in the claims can be found, for example, at page 40, lines 18-19. Support for the stringent hybridization conditions may be found, for example, at page 14, lines 10-14. Support for the recitation of sequence relatedness such as those recited in claim 52 can be found in the specification, for example, at page 44, lines 1-25.

Closing Remarks

Entry of this Preliminary Amendment prior to calculation of fees and Examination and Allowance of the pending claims is respectfully requested.

Respectfully submitted,

Arthur E. Jackson
Registration No. 34,354
Allen Bloom
Registration No. 29,135
Attorneys for Applicant

DECHERT PRICE & RHOADS
Princeton Pike Corporate Center
PO Box 5218
Princeton, New Jersey 08543-5218
Allen Bloom (609) 620-3214
Arthur E. Jackson (609) 620-3254
Fax: (609) 620-3259
Attn: Allen Bloom, Esq.
(609) 620-3214

SEQUENCE LISTING

<110> Thonnard, Joelle

<120> Novel Compounds

<130> BM45330

<150> PCT/EP99/05989
<151> 1999-08-13

<150> GB 9818004.5
<151> 1998-08-18

<160> 9

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 2769
<212> DNA
<213> Bacteria

<400> 1
atgagatctt ctttccggtt gaagccgatt tgtttttatc ttatgggtgt tatgctatat 60
catcatagtt atgccgaaga tgcagggcgc gcgggcagcg aggcgcagat acaggttttg 120
gaagatgtgc acgtcaaggc gaagcgcgta ccgaaagaca aaaaagtgtt taccgatgcg 180
cgtgccgtat cgacccgtca ggatatattc aaatccagcg aaaacctcga caacatcgta 240
cgcagcatcc ccggtgcgtt tacacagcaa gataaaagct cgggcattgt gtctttgaat 300
attcgcggcg acagcgggtt cgggcgggtc aatacgatgg tggacggcat cacgcagacc 360
ttttattcga cttctaccga tgcgggcagg gcaggcggtt catctcaatt cggtgcatct 420
gtcgacagca attttattgc cggactggat gtcgtcaaag gcagcttcag cggctcggca 480
ggcatcaaca gccttgccgg ttcggcgaat ctgcggactt taggcgtgga tgacgtcgtt 540
cagggcaata atacctacgg cctgctgcta aaaggtctga ccggcaccaa ttcaaccaaa 600
ggtaatgcga tggcggcgat aggtgcgcgc aaatggctgg aaagcggagc atctgtcggt 660
gtgctttacg ggcacagcag gcgcacgtgg gcgcaaaatt accgcgtggg cggcggcggg 720
cagcacatcg gaaattttgg cgcggaatat ctggaacggc gcaaacagcg atattttgta 780
caagaaggcg ggttgaaatt caattccaac agcggaaaat gggagcggga tttccaaagg 840
ccgtactgga aaaccaagtg gtatcaaaaa tacaatgacc cccaagaact gcaaaaatac 900
atcgaaggtc atgacaaaag ctggcgggaa aacctggcgc cgcaatacga catcaccccc 960
atcgatccgt ccagcctgaa gcagcagtcg gcaggcaatc tgtttaaatt ggaatacgac 1020
ggcgtattca ataaatacac ggcgcaattt cgcgatttaa acaccaaaat cggcagccgc 1080
aaaatcatca accgcaatta tcaattcaat tacggtttat ctttaaactc atatgccaac 1140
ctcaatctga ccgcagccta caattcgggc aggcagaaat atccgaaagg gtcgaagttt 1200
acaggctggg ggcttttaaa agattttgaa acctacaaca acgcgaaaat cctcgacctc 1260
aacaacaccg ccaccttccg gctgccccgc gaaaccgagt tgcaaaccac tttgggcttc 1320
aattatttcc acaacgaata cggcaaaaac cgctttcctg aagaattggg gctgtttttc 1380
gacggtccgg atcaggacaa cgggctttat tcctatttgg ggcggtttaa gggcgataaa 1440
gggctgctgc cccaaaaatc aaccatcgtc caaccggccg gcagccaata tttcaacacg 1500
ttctacttcg atgccgcgct caaaaaagac atttaccgct taaactacag caccaatacc 1560
gtcggctacc gtttcggcgg cgaatatacg ggctattacg gctcggatga cgaatttaag 1620
cgggcattcg gagaaaactc gccgacatac aagaaacatt gcaaccagag ctgcggaatt 1680
tatgaacccg tattgaaaaa atacggcaaa aagcgcgcca acaaccattc ggtcagcatt 1740
agtgcggact tcggcgatta tttcatgccg ttcgccagct attcgcgcac acaccgtatg 1800
cccaacatcc aagaaatgta tttttcccaa atcggcgact ccggcgttca caccgcctta 1860
aaaccagagc gcgcaaacac ttggcaattt ggcttcaata cctataaaaa aggattgtta 1920
aaacaagatg atacattagg attaaaactg gtcggctacc gcagccgcat cgacaactac 1980
atccacaacg tttacgggaa atggtgggat ttgaacggga atattccgag ctgggtcagc 2040
agcaccgggc ttgcctacac catccaacac cgcaatttca aagacaaagt acacaaacac 2100
ggttttgagt tggagctgaa ttacgattat gggcgttttt tcaccaacct ttcttacgcc 2160

tatcaaaaaa gcacgcaacc gaccaacttc agcgatgcga gcgaatcgcc caacaatgcg 2220
tccaaagaag accaactcaa acaaggttat gggttgagca gggtttccgc cctgccgcga 2280
gattacggac gtttggaagt cggtacgcgc tggttgggca acaaactgac tttgggcggc 2340
gcgatgcgct atttcggcaa gagcatccgc gcgacggctg aagaacgcta tatcgacggc 2400
accaacgggg gaaataccag caatgtccgg caactgggca agcgttccat caaacaaacc 2460
gaaacccttg cccgccagcc tttgattttt gatttttacg ccgcttacga gccgaagaaa 2520
aaccttattt tccgcgccga agtcaaaaat ctgttcgaca ggcgttatat cgatccgctc 2580
gatgcgggca atgatgcggc aacgcagcgt tattacagtt cgttcgaccc gaaagacaag 2640
gacgaagaag taacgtgtaa tgctgataaa acgttgtgca acggcaaata cggcggcaca 2700
agcaaaagcg tattgaccaa ttttgcacgc ggacgcacct tttgataac gatgagctac 2760
aagttttaa 2769

<210> 2
<211> 922
<212> PRT
<213> Bacteria

<400> 2
Met Arg Ser Ser Phe Arg Leu Lys Pro Ile Cys Phe Tyr Leu Met Gly
1 5 10 15
Val Met Leu Tyr His His Ser Tyr Ala Glu Asp Ala Gly Arg Ala Gly
20 25 30
Ser Glu Ala Gln Ile Gln Val Leu Glu Asp Val His Val Lys Ala Lys
35 40 45
Arg Val Pro Lys Asp Lys Lys Val Phe Thr Asp Ala Arg Ala Val Ser
50 55 60
Thr Arg Gln Asp Ile Phe Lys Ser Ser Glu Asn Leu Asp Asn Ile Val
65 70 75 80
Arg Ser Ile Pro Gly Ala Phe Thr Gln Gln Asp Lys Ser Ser Gly Ile
85 90 95
Val Ser Leu Asn Ile Arg Gly Asp Ser Gly Phe Gly Arg Val Asn Thr
100 105 110
Met Val Asp Gly Ile Thr Gln Thr Phe Tyr Ser Thr Ser Thr Asp Ala
115 120 125
Gly Arg Ala Gly Gly Ser Ser Gln Phe Gly Ala Ser Val Asp Ser Asn
130 135 140
Phe Ile Ala Gly Leu Asp Val Val Lys Gly Ser Phe Ser Gly Ser Ala
145 150 155 160
Gly Ile Asn Ser Leu Ala Gly Ser Ala Asn Leu Arg Thr Leu Gly Val
165 170 175
Asp Asp Val Val Gln Gly Asn Asn Thr Tyr Gly Leu Leu Leu Lys Gly
180 185 190
Leu Thr Gly Thr Asn Ser Thr Lys Gly Asn Ala Met Ala Ala Ile Gly
195 200 205
Ala Arg Lys Trp Leu Glu Ser Gly Ala Ser Val Gly Val Leu Tyr Gly
210 215 220
His Ser Arg Arg Thr Trp Ala Gln Asn Tyr Arg Val Gly Gly Gly Gly
225 230 235 240
Gln His Ile Gly Asn Phe Gly Ala Glu Tyr Leu Glu Arg Arg Lys Gln
245 250 255
Arg Tyr Phe Val Gln Glu Gly Gly Leu Lys Phe Asn Ser Asn Ser Gly
260 265 270
Lys Trp Glu Arg Asp Phe Gln Arg Pro Tyr Trp Lys Thr Lys Trp Tyr
275 280 285
Gln Lys Tyr Asn Asp Pro Gln Glu Leu Gln Lys Tyr Ile Glu Gly His
290 295 300
Asp Lys Ser Trp Arg Glu Asn Leu Ala Pro Gln Tyr Asp Ile Thr Pro
305 310 315 320
Ile Asp Pro Ser Ser Leu Lys Gln Gln Ser Ala Gly Asn Leu Phe Lys
325 330 335
Leu Glu Tyr Asp Gly Val Phe Asn Lys Tyr Thr Ala Gln Phe Arg Asp
340 345 350

```
Leu Asn Thr Lys Ile Gly Ser Arg Lys Ile Ile Asn Arg Asn Tyr Gln
            355                 360                 365
Phe Asn Tyr Gly Leu Ser Leu Asn Ser Tyr Ala Asn Leu Asn Leu Thr
    370                 375                 380
Ala Ala Tyr Asn Ser Gly Arg Gln Lys Tyr Pro Lys Gly Ser Lys Phe
385                 390                 395                 400
Thr Gly Trp Gly Leu Leu Lys Asp Phe Glu Thr Tyr Asn Asn Ala Lys
                405                 410                 415
Ile Leu Asp Leu Asn Asn Thr Ala Thr Phe Arg Leu Pro Arg Glu Thr
            420                 425                 430
Glu Leu Gln Thr Thr Leu Gly Phe Asn Tyr Phe His Asn Glu Tyr Gly
            435                 440                 445
Lys Asn Arg Phe Pro Glu Glu Leu Gly Leu Phe Phe Asp Gly Pro Asp
    450                 455                 460
Gln Asp Asn Gly Leu Tyr Ser Tyr Leu Gly Arg Phe Lys Gly Asp Lys
465                 470                 475                 480
Gly Leu Leu Pro Gln Lys Ser Thr Ile Val Gln Pro Ala Gly Ser Gln
                485                 490                 495
Tyr Phe Asn Thr Phe Tyr Phe Asp Ala Ala Leu Lys Lys Asp Ile Tyr
            500                 505                 510
Arg Leu Asn Tyr Ser Thr Asn Thr Val Gly Tyr Arg Phe Gly Gly Glu
            515                 520                 525
Tyr Thr Gly Tyr Tyr Gly Ser Asp Asp Glu Phe Lys Arg Ala Phe Gly
    530                 535                 540
Glu Asn Ser Pro Thr Tyr Lys Lys His Cys Asn Gln Ser Cys Gly Ile
545                 550                 555                 560
Tyr Glu Pro Val Leu Lys Lys Tyr Gly Lys Lys Arg Ala Asn Asn His
                565                 570                 575
Ser Val Ser Ile Ser Ala Asp Phe Gly Asp Tyr Phe Met Pro Phe Ala
            580                 585                 590
Ser Tyr Ser Arg Thr His Arg Met Pro Asn Ile Gln Glu Met Tyr Phe
            595                 600                 605
Ser Gln Ile Gly Asp Ser Gly Val His Thr Ala Leu Lys Pro Glu Arg
    610                 615                 620
Ala Asn Thr Trp Gln Phe Gly Phe Asn Thr Tyr Lys Lys Gly Leu Leu
625                 630                 635                 640
Lys Gln Asp Asp Thr Leu Gly Leu Lys Leu Val Gly Tyr Arg Ser Arg
                645                 650                 655
Ile Asp Asn Tyr Ile His Asn Val Tyr Gly Lys Trp Trp Asp Leu Asn
            660                 665                 670
Gly Asn Ile Pro Ser Trp Val Ser Ser Thr Gly Leu Ala Tyr Thr Ile
            675                 680                 685
Gln His Arg Asn Phe Lys Asp Lys Val His Lys His Gly Phe Glu Leu
    690                 695                 700
Glu Leu Asn Tyr Asp Tyr Gly Arg Phe Phe Thr Asn Leu Ser Tyr Ala
705                 710                 715                 720
Tyr Gln Lys Ser Thr Gln Pro Thr Asn Phe Ser Asp Ala Ser Glu Ser
                725                 730                 735
Pro Asn Asn Ala Ser Lys Glu Asp Gln Leu Lys Gln Gly Tyr Gly Leu
            740                 745                 750
Ser Arg Val Ser Ala Leu Pro Arg Asp Tyr Gly Arg Leu Glu Val Gly
            755                 760                 765
Thr Arg Trp Leu Gly Asn Lys Leu Thr Leu Gly Gly Ala Met Arg Tyr
    770                 775                 780
Phe Gly Lys Ser Ile Arg Ala Thr Ala Glu Glu Arg Tyr Ile Asp Gly
785                 790                 795                 800
Thr Asn Gly Gly Asn Thr Ser Asn Val Arg Gln Leu Gly Lys Arg Ser
                805                 810                 815
Ile Lys Gln Thr Glu Thr Leu Ala Arg Gln Pro Leu Ile Phe Asp Phe
            820                 825                 830
Tyr Ala Ala Tyr Glu Pro Lys Lys Asn Leu Ile Phe Arg Ala Glu Val
            835                 840                 845
```

```
Lys Asn Leu Phe Asp Arg Arg Tyr Ile Asp Pro Leu Asp Ala Gly Asn
    850                 855                 860
Asp Ala Ala Thr Gln Arg Tyr Tyr Ser Ser Phe Asp Pro Lys Asp Lys
865                 870                 875                 880
Asp Glu Glu Val Thr Cys Asn Ala Asp Lys Thr Leu Cys Asn Gly Lys
                885                 890                 895
Tyr Gly Gly Thr Ser Lys Ser Val Leu Thr Asn Phe Ala Arg Gly Arg
            900                 905                 910
Thr Phe Leu Ile Thr Met Ser Tyr Lys Phe
        915                 920

<210> 3
<211> 2769
<212> DNA
<213> Bacteria

<400> 3
atgagatctt ctttccggtt gaagccgatt tgtttttatc ttatgggtgt tatgctatat      60
catcatagtt atgccgaaga tgcagggcgc gcgggcagcg aggcgcagat acaggttttg     120
gaagatgtgc acgtcaaggc gaagcgcgta ccgaaagaca aaaaagtgtt taccgatgcg     180
cgtgccgtat cgacccgtca ggatatattc aaatccagcg aaaacctcga caacatcgta     240
cgcagcatcc ccggtgcgtt tacacagcaa gataaaagct cgggcattgt gtctttgaat     300
attcgcggcg acagcgggtt cgggcgggtc aatacgatgg tggacggcat cacgcagacc     360
ttttattcga cttctaccga tgcgggcagg gcaggcggtt catctcaatt cggtgcatct     420
gtcgacagca attttattgc cggactggat gtcgtcaaag gcagcttcag cggctcggca     480
ggcatcaaca gcctttgccgg ttcggcgaat ctgcggactt taggcgtgga tgacgtcgtt     540
cagggcaata atacctacgg cctgctgcta aaaggtctga ccggcaccaa ttcaaccaaa     600
ggtaatgcga tggcggcgat aggtgcgcgc aaatggctgg aaagcggagc atctgtcggt     660
gtgctttacg ggcacagcag gcgcagcgtg gcgcaaaatt accgcgtggg cggcggcggg     720
cagcacatcg gaaattttgg cgcggaatat ctggaacggc gcaagcagcg atattttgta     780
caagaaggcg ggttgaaatt caattccaac agcggaaaat gggagcggga tttccaaagg     840
ccgtactgga aaaccaagtg gtatcaaaaa tacaatgacc cccaagaact gcaaaaatac     900
atcgaaggtc atgacaaaag ctggcgggaa aacctggcgc cgcaatacga catcaccccc     960
atcgatccgt ccagcctgaa gcagcagtcg gcaggcaatc tgtttaaatt ggaatacgac    1020
ggcgtattca ataaatacac ggcgcaattt cgcgatttaa acaccaaaat cggcagccgc    1080
aaaatcatca accgcaatta tcaattcaat tacggtttat ctttaaactc atatgccaac    1140
ctcaatctga ccgcagccta caattcgggc aggcagaaat atccgaaagg gtcgaagttt    1200
acaggctggg ggcttttaaa agattttgaa acctacaaca acgcgaaaat cctcgacctc    1260
aacaacaccg ccaccttccg gctgccccgc gaaaccgagt tgcaaaccac tttggcgctc    1320
aattatttcc acaacgaata cggcaaaaac cgctttcctg aagaattggg gctgtttttc    1380
gacggtccgg atcaggacaa cgggctttat tcctatttgg ggcggttaa gggcgataaa    1440
gggctgctgc cccaaaaatc aaccatcgtc caaccggccg gcagccaata tttcaacacg    1500
ttctacttcg atgccgcgct caaaaaagac atttaccgct taaactacag caccaatacc    1560
gtcggctacc gtttcggcgg cgaatatacg ggctattacg gctcggatga cgaatttaag    1620
cgggcattcg gagaaaactc gccgacatac aagaaacatt gcaaccagag ctgcggaatt    1680
tatgaacccg tattgaaaaa atacggcaaa aagcgcgcca acaaccattc ggtcagcatt    1740
agtgcggact tcggcgatta tttcatgccg ttcgccagct attcgcgcac acaccgtatg    1800
cccaacatcc aagaaatgta tcttttcccaa atcggcgact ccggcgttca caccgcctta    1860
aaaccagagc gcgcaaacac ttggcaattt ggcttcaata cctataaaaa aggattgtta    1920
aaacaagatg atacattagg attaaaactg gtcggctacc gcagccgcat cgacaactac    1980
atccacaacg tttacgggaa atggtgggat ttgaacggga atattccgag ctgggtcagc    2040
agcaccgggc ttgcctacac catccaacac cgcaatttca aagacaaagt acacaaaacc    2100
ggttttgagt tggagctgaa ttacgattat gggcgttttt tcaccaacct ttcttacgcc    2160
tatcaaaaaa gcacgcaacc gaccaacttc agcgatgcga gcgaatcgcc caacaatgcg    2220
tccaaagaag accaactcaa acaaggttat gggttgagca gggtttccgc cctgccgcga    2280
gattacggac gtttggaagt cggtacgcgc tggttgggca acaaactgac tttggcggc     2340
gcgatgcgct atttcggcaa gagcatccgc gcgacggctg aagaacgcta tatcgacggc    2400
accaacgggg gaaataccag caatgtccgg caactgggca agcgttccat caaacaaacc    2460
gaaacccttg cccgccagcc tttgattttt gatttttacg ccgcttacga gccgaagaaa    2520
aaccttattt tccgcgccga agtcaaaaat ctgttcgaca ggcgttatat cgatccgctc    2580
gatgcgggca atgatgcggc aacgcagcgt tattacagtt cgttcgaccc gaaagacaag    2640
gacgaagaag taacgtgtaa tgctgataaa acgttgtgca acggcaaata cggcggcaca    2700
```

```
agcaaaagcg tattgaccaa ttttgcacgc ggacgcacct ttttgataac gatgagctac      2760
aagttttaa                                                              2769

<210> 4
<211> 922
<212> PRT
<213> Bacteria

<400> 4
Met Arg Ser Ser Phe Arg Leu Lys Pro Ile Cys Phe Tyr Leu Met Gly
1               5                   10                  15
Val Met Leu Tyr His His Ser Tyr Ala Glu Asp Ala Gly Arg Ala Gly
            20                  25                  30
Ser Glu Ala Gln Ile Gln Val Leu Glu Asp Val His Val Lys Ala Lys
        35                  40                  45
Arg Val Pro Lys Asp Lys Lys Val Phe Thr Asp Ala Arg Ala Val Ser
    50                  55                  60
Thr Arg Gln Asp Ile Phe Lys Ser Ser Glu Asn Leu Asp Asn Ile Val
65                  70                  75                  80
Arg Ser Ile Pro Gly Ala Phe Thr Gln Gln Asp Lys Ser Ser Gly Ile
                85                  90                  95
Val Ser Leu Asn Ile Arg Gly Asp Ser Gly Phe Gly Arg Val Asn Thr
            100                 105                 110
Met Val Asp Gly Ile Thr Gln Thr Phe Tyr Ser Thr Ser Thr Asp Ala
        115                 120                 125
Gly Arg Ala Gly Gly Ser Ser Gln Phe Gly Ala Ser Val Asp Ser Asn
    130                 135                 140
Phe Ile Ala Gly Leu Asp Val Val Lys Gly Ser Phe Ser Gly Ser Ala
145                 150                 155                 160
Gly Ile Asn Ser Leu Ala Gly Ser Ala Asn Leu Arg Thr Leu Gly Val
                165                 170                 175
Asp Asp Val Val Gln Gly Asn Asn Thr Tyr Gly Leu Leu Leu Lys Gly
            180                 185                 190
Leu Thr Gly Thr Asn Ser Thr Lys Gly Asn Ala Met Ala Ala Ile Gly
        195                 200                 205
Ala Arg Lys Trp Leu Glu Ser Gly Ala Ser Val Gly Val Leu Tyr Gly
    210                 215                 220
His Ser Arg Arg Ser Val Ala Gln Asn Tyr Arg Val Gly Gly Gly Gly
225                 230                 235                 240
Gln His Ile Gly Asn Phe Gly Ala Glu Tyr Leu Glu Arg Arg Lys Gln
                245                 250                 255
Arg Tyr Phe Val Gln Glu Gly Gly Leu Lys Phe Asn Ser Asn Ser Gly
            260                 265                 270
Lys Trp Glu Arg Asp Phe Gln Arg Pro Tyr Trp Lys Thr Lys Trp Tyr
        275                 280                 285
Gln Lys Tyr Asn Asp Pro Gln Glu Leu Gln Lys Tyr Ile Glu Gly His
    290                 295                 300
Asp Lys Ser Trp Arg Glu Asn Leu Ala Pro Gln Tyr Asp Ile Thr Pro
305                 310                 315                 320
Ile Asp Pro Ser Ser Leu Lys Gln Gln Ser Ala Gly Asn Leu Phe Lys
                325                 330                 335
Leu Glu Tyr Asp Gly Val Phe Asn Lys Tyr Thr Ala Gln Phe Arg Asp
            340                 345                 350
Leu Asn Thr Lys Ile Gly Ser Arg Lys Ile Ile Asn Arg Asn Tyr Gln
        355                 360                 365
Phe Asn Tyr Gly Leu Ser Leu Asn Ser Tyr Ala Asn Leu Asn Leu Thr
    370                 375                 380
Ala Ala Tyr Asn Ser Gly Arg Gln Lys Tyr Pro Lys Gly Ser Lys Phe
385                 390                 395                 400
Thr Gly Trp Gly Leu Leu Lys Asp Phe Glu Thr Tyr Asn Asn Ala Lys
                405                 410                 415
Ile Leu Asp Leu Asn Asn Thr Ala Thr Phe Arg Leu Pro Arg Glu Thr
```

```
            420                 425                 430
Glu Leu Gln Thr Thr Leu Gly Phe Asn Tyr Phe His Asn Glu Tyr Gly
        435                 440                 445
Lys Asn Arg Phe Pro Glu Glu Leu Gly Leu Phe Phe Asp Gly Pro Asp
    450                 455                 460
Gln Asp Asn Gly Leu Tyr Ser Tyr Leu Gly Arg Phe Lys Gly Asp Lys
465                 470                 475                 480
Gly Leu Leu Pro Gln Lys Ser Thr Ile Val Gln Pro Ala Gly Ser Gln
                485                 490                 495
Tyr Phe Asn Thr Phe Tyr Phe Asp Ala Ala Leu Lys Lys Asp Ile Tyr
            500                 505                 510
Arg Leu Asn Tyr Ser Thr Asn Thr Val Gly Tyr Arg Phe Gly Gly Glu
        515                 520                 525
Tyr Thr Gly Tyr Tyr Gly Ser Asp Asp Glu Phe Lys Arg Ala Phe Gly
    530                 535                 540
Glu Asn Ser Pro Thr Tyr Lys Lys His Cys Asn Gln Ser Cys Gly Ile
545                 550                 555                 560
Tyr Glu Pro Val Leu Lys Lys Tyr Gly Lys Lys Arg Ala Asn Asn His
                565                 570                 575
Ser Val Ser Ile Ser Ala Asp Phe Gly Asp Tyr Phe Met Pro Phe Ala
            580                 585                 590
Ser Tyr Ser Arg Thr His Arg Met Pro Asn Ile Gln Glu Met Tyr Phe
        595                 600                 605
Ser Gln Ile Gly Asp Ser Gly Val His Thr Ala Leu Lys Pro Glu Arg
    610                 615                 620
Ala Asn Thr Trp Gln Phe Gly Phe Asn Thr Tyr Lys Lys Gly Leu Leu
625                 630                 635                 640
Lys Gln Asp Asp Thr Leu Gly Leu Lys Leu Val Gly Tyr Arg Ser Arg
                645                 650                 655
Ile Asp Asn Tyr Ile His Asn Val Tyr Gly Lys Trp Trp Asp Leu Asn
            660                 665                 670
Gly Asn Ile Pro Ser Trp Val Ser Ser Thr Gly Leu Ala Tyr Thr Ile
        675                 680                 685
Gln His Arg Asn Phe Lys Asp Lys Val His Lys His Gly Phe Glu Leu
    690                 695                 700
Glu Leu Asn Tyr Asp Tyr Gly Arg Phe Phe Thr Asn Leu Ser Tyr Ala
705                 710                 715                 720
Tyr Gln Lys Ser Thr Gln Pro Thr Asn Phe Ser Asp Ala Ser Glu Ser
                725                 730                 735
Pro Asn Asn Ala Ser Lys Glu Asp Gln Leu Lys Gln Gly Tyr Gly Leu
            740                 745                 750
Ser Arg Val Ser Ala Leu Pro Arg Asp Tyr Gly Arg Leu Glu Val Gly
        755                 760                 765
Thr Arg Trp Leu Gly Asn Lys Leu Thr Leu Gly Gly Ala Met Arg Tyr
    770                 775                 780
Phe Gly Lys Ser Ile Arg Ala Thr Ala Glu Glu Arg Tyr Ile Asp Gly
785                 790                 795                 800
Thr Asn Gly Gly Asn Thr Ser Asn Val Arg Gln Leu Gly Lys Arg Ser
                805                 810                 815
Ile Lys Gln Thr Glu Thr Leu Ala Arg Gln Pro Leu Ile Phe Asp Phe
            820                 825                 830
Tyr Ala Ala Tyr Glu Pro Lys Lys Asn Leu Ile Phe Arg Ala Glu Val
        835                 840                 845
Lys Asn Leu Phe Asp Arg Arg Tyr Ile Asp Pro Leu Asp Ala Gly Asn
    850                 855                 860
Asp Ala Ala Thr Gln Arg Tyr Tyr Ser Ser Phe Asp Pro Lys Asp Lys
865                 870                 875                 880
Asp Glu Glu Val Thr Cys Asn Ala Asp Lys Thr Leu Cys Asn Gly Lys
                885                 890                 895
Tyr Gly Gly Thr Ser Lys Ser Val Leu Thr Asn Phe Ala Arg Gly Arg
            900                 905                 910
Thr Phe Leu Ile Thr Met Ser Tyr Lys Phe
```

915   920

<210> 5
<211> 2766
<212> DNA
<213> Bacteria

<400> 5
```
atgagatctt ctttccggtt gaagccgatt tgtttttacc ttatgggtgt tacgctatat      60
cattatagtt atgccgaaga tgcagggcgc gcgggcagcg aggcgcagat acaggttttg     120
gaagatgtgc acgtcaaggc gaagcgcgta ccgaaagaca aaaaagtgtt taccgatgcg     180
cgtgccgtat cgacccgtca ggatatattc aaatccagcg aaaacctcga caacatcgta     240
cgcagcatcc ccggtgcgtt tacacagcaa gataaaagct cgggcattgt gtctttgaat     300
attcgcggcg acagcgggtt cgggcgggtc aatacgatgg tggacggcat cacgcagacc     360
ttttattcga cttctaccga tgcgggcagg gcaggcggtt catctcaatt cggtgcatct     420
gtcgacagca attttattgc cggactggat gtcgtcaaag gcagcttcag cggctcggca     480
ggcatcaaca gccttgccgg ttcggcgaat ctgcggactt taggcgtgga tgacgtcgtt     540
cagggcaata atacctacgg cctgctgcta aaaggtctga ccggcaccaa ttcaaccaaa     600
ggtaatgcga tggcggcgat aggtgcgcgc aaatggctgg aaagcggagc atctgtcggt     660
gtgctttacg ggcacagcag gcgcagcgtg gcgcaaaatt accgcgtggg cggcggcggg     720
cagcacatcg gaaattttgg cgcggaatat ttggaacggc gcaagcagcg atattttgta     780
caagagggtg ctttgaaatt caattccgac agcggaaaat gggagcggga tttacaaagg     840
caacagtgga aatacaagcc gtataaaaat tacaacaacc aagaactaca aaaatacatc     900
gaaggtcatg acaaaagctg gcgggaaaac ctggcgccgc aatacgacat cacccccatc     960
gatccgtcca gcctgaagca gcagtcggca ggcaatctgt ttaaattgga atacgacggc    1020
gtattcaata aatacacggc gcaatttcgc gatttaaaca ccaaaatcgg cagccgcaaa    1080
atcatcaacc gcaattatca gttcaattac ggtttgtctt tgaacccgta taccaacctc    1140
aatctgaccg cagcctacaa ttcgggcagg cagaaatatc cgaaagggtc gaagtttaca    1200
ggctgggggc ttttaaagga ttttgaaacc tacaacaacg cgaaaatcct cgacctcaac    1260
aacaccgcca ccttccggct gccccgcgaa accgagttgc aaaccacttt gggcttcaat    1320
tatttccaca acgaatacgg caaaaaccgc tttcctgaag aattggggct gtttttcgac    1380
ggtcctgatc aggacaacgg gctttattcc tatttggggc ggtttaaggg cgataaaggg    1440
ctgctgcccc aaaaatcaac cattgtccaa ccggccggca gccaatattt caacacgttc    1500
tacttcgatg ccgcgctcaa aaaagacatt taccgcttaa actacagcac caataccgtc    1560
ggctaccgtt tcggcggcga atatacgggc tattacggct cggatgacga atttaagcgg    1620
gcattcggag aaaactcgcc gacatacaag aaacattgca accggagctg cgggatttat    1680
gaacccgtat tgaaaaaata cggcaaaaag cgcgccaaca accattcggt cagcattagt    1740
gcggacttcg gcgattattt catgccgttc gccagctatt cgcgcacaca ccgtatgccc    1800
aacatccaag aaatgtattt ttcccaaatc ggcgactccg gcgttcacac cgccttaaaa    1860
ccagagcgcg caaacacttg gcaatttggc ttcaatacct ataaaaaagg attgttaaaa    1920
caagatgata cattaggatt aaaactggtc ggctaccgca gccgcatcga caactacatc    1980
cacaacgttt acgggaaatg gtgggatttg aacggggata ttccgagctg ggtcagcagc    2040
accgggcttg cctacaccat ccaacatcgc aatttcaaag acaaagtgca caaacacggt    2100
tttgagttgg agctgaatta cgattatggg cgtttttttca ccaacctttc ttacgcctat    2160
caaaaaagca cgcaaccgac caacttcagc gatgcgagcg aatcgcccaa caatgcgtcc    2220
aaagaagacc aactcaaaca aggttatggg ttgagcaggg tttccgccct gccgcgagat    2280
tacggacgtt tggaagtcgg tacgcgctgg ttgggcaaca aactgacttt gggcggcgcg    2340
atgcgctatt tcggcaagag catccgcgcg acggctgaag aacgctatat cgacggcacc    2400
aacgggggaa ataccagcaa tttccggcaa ctgggcaagc gttccatcaa acaaaccgaa    2460
actcttgccc gccagccttt gatttttgat ttttacgccg cttacgagcc gaagaaaaac    2520
cttattttcc gcgccgaagt caaaaatctg ttcgacaggc gttatatcga tccgctcgat    2580
gcgggcaatg atgcggcaac gcagcgttat tacagctcgt tcgacccgaa agacaaggac    2640
gaagacgtaa cgtgtaatgc tgataaaacg ttgtgcaacg gcaaatacgg cggcacaagc    2700
aaaagcgtat tgaccaattt tgcacgcgga cgcacctttt tgataacgat gagctacaag    2760
ttttaa                                                               2766
```

<210> 6
<211> 921
<212> PRT
<213> Bacteria

<400> 6

```
Met Arg Ser Ser Phe Arg Leu Lys Pro Ile Cys Phe Tyr Leu Met Gly
 1               5                  10                  15
Val Thr Leu Tyr His Tyr Ser Tyr Ala Glu Asp Ala Gly Arg Ala Gly
            20                  25                  30
Ser Glu Ala Gln Ile Gln Val Leu Glu Asp Val His Val Lys Ala Lys
        35                  40                  45
Arg Val Pro Lys Asp Lys Lys Val Phe Thr Asp Ala Arg Ala Val Ser
    50                  55                  60
Thr Arg Gln Asp Ile Phe Lys Ser Ser Glu Asn Leu Asp Asn Ile Val
65                  70                  75                  80
Arg Ser Ile Pro Gly Ala Phe Thr Gln Gln Asp Lys Ser Ser Gly Ile
                85                  90                  95
Val Ser Leu Asn Ile Arg Gly Asp Ser Gly Phe Gly Arg Val Asn Thr
            100                 105                 110
Met Val Asp Gly Ile Thr Gln Thr Phe Tyr Ser Thr Ser Thr Asp Ala
        115                 120                 125
Gly Arg Ala Gly Gly Ser Ser Gln Phe Gly Ala Ser Val Asp Ser Asn
    130                 135                 140
Phe Ile Ala Gly Leu Asp Val Val Lys Gly Ser Phe Ser Gly Ser Ala
145                 150                 155                 160
Gly Ile Asn Ser Leu Ala Gly Ser Ala Asn Leu Arg Thr Leu Gly Val
                165                 170                 175
Asp Asp Val Val Gln Gly Asn Asn Thr Tyr Gly Leu Leu Leu Lys Gly
            180                 185                 190
Leu Thr Gly Thr Asn Ser Thr Lys Gly Asn Ala Met Ala Ala Ile Gly
        195                 200                 205
Ala Arg Lys Trp Leu Glu Ser Gly Ala Ser Val Gly Val Leu Tyr Gly
    210                 215                 220
His Ser Arg Arg Ser Val Ala Gln Asn Tyr Arg Val Gly Gly Gly Gly
225                 230                 235                 240
Gln His Ile Gly Asn Phe Gly Ala Glu Tyr Leu Glu Arg Arg Lys Gln
                245                 250                 255
Arg Tyr Phe Val Gln Glu Gly Ala Leu Lys Phe Asn Ser Asp Ser Gly
            260                 265                 270
Lys Trp Glu Arg Asp Leu Gln Arg Gln Gln Trp Lys Tyr Lys Pro Tyr
        275                 280                 285
Lys Asn Tyr Asn Asn Gln Glu Leu Gln Lys Tyr Ile Glu Gly His Asp
    290                 295                 300
Lys Ser Trp Arg Glu Asn Leu Ala Pro Gln Tyr Asp Ile Thr Pro Ile
305                 310                 315                 320
Asp Pro Ser Ser Leu Lys Gln Gln Ser Ala Gly Asn Leu Phe Lys Leu
                325                 330                 335
Glu Tyr Asp Gly Val Phe Asn Lys Tyr Thr Ala Gln Phe Arg Asp Leu
            340                 345                 350
Asn Thr Lys Ile Gly Ser Arg Lys Ile Ile Asn Arg Asn Tyr Gln Phe
        355                 360                 365
Asn Tyr Gly Leu Ser Leu Asn Pro Tyr Thr Asn Leu Asn Leu Thr Ala
    370                 375                 380
Ala Tyr Asn Ser Gly Arg Gln Lys Tyr Pro Lys Gly Ser Lys Phe Thr
385                 390                 395                 400
Gly Trp Gly Leu Leu Lys Asp Phe Glu Thr Tyr Asn Asn Ala Lys Ile
                405                 410                 415
Leu Asp Leu Asn Asn Thr Ala Thr Phe Arg Leu Pro Arg Glu Thr Glu
            420                 425                 430
Leu Gln Thr Thr Leu Gly Phe Asn Tyr Phe His Asn Glu Tyr Gly Lys
        435                 440                 445
Asn Arg Phe Pro Glu Glu Leu Gly Leu Phe Phe Asp Gly Pro Asp Gln
    450                 455                 460
Asp Asn Gly Leu Tyr Ser Tyr Leu Gly Arg Phe Lys Gly Asp Lys Gly
465                 470                 475                 480
Leu Leu Pro Gln Lys Ser Thr Ile Val Gln Pro Ala Gly Ser Gln Tyr
                485                 490                 495
```

```
Phe Asn Thr Phe Tyr Phe Asp Ala Ala Leu Lys Lys Asp Ile Tyr Arg
            500                 505                 510
Leu Asn Tyr Ser Thr Asn Thr Val Gly Tyr Arg Phe Gly Gly Glu Tyr
        515                 520                 525
Thr Gly Tyr Tyr Gly Ser Asp Asp Glu Phe Lys Arg Ala Phe Gly Glu
    530                 535                 540
Asn Ser Pro Thr Tyr Lys Lys His Cys Asn Arg Ser Cys Gly Ile Tyr
545                 550                 555                 560
Glu Pro Val Leu Lys Lys Tyr Gly Lys Lys Arg Ala Asn Asn His Ser
                565                 570                 575
Val Ser Ile Ser Ala Asp Phe Gly Asp Tyr Phe Met Pro Phe Ala Ser
            580                 585                 590
Tyr Ser Arg Thr His Arg Met Pro Asn Ile Gln Glu Met Tyr Phe Ser
        595                 600                 605
Gln Ile Gly Asp Ser Gly Val His Thr Ala Leu Lys Pro Glu Arg Ala
    610                 615                 620
Asn Thr Trp Gln Phe Gly Phe Asn Thr Tyr Lys Lys Gly Leu Leu Lys
625                 630                 635                 640
Gln Asp Asp Thr Leu Gly Leu Lys Leu Val Gly Tyr Arg Ser Arg Ile
                645                 650                 655
Asp Asn Tyr Ile His Asn Val Tyr Gly Lys Trp Trp Asp Leu Asn Gly
            660                 665                 670
Asp Ile Pro Ser Trp Val Ser Ser Thr Gly Leu Ala Tyr Thr Ile Gln
        675                 680                 685
His Arg Asn Phe Lys Asp Lys Val His Lys His Gly Phe Glu Leu Glu
    690                 695                 700
Leu Asn Tyr Asp Tyr Gly Arg Phe Phe Thr Asn Leu Ser Tyr Ala Tyr
705                 710                 715                 720
Gln Lys Ser Thr Gln Pro Thr Asn Phe Ser Asp Ala Ser Glu Ser Pro
                725                 730                 735
Asn Asn Ala Ser Lys Glu Asp Gln Leu Lys Gln Gly Tyr Gly Leu Ser
            740                 745                 750
Arg Val Ser Ala Leu Pro Arg Asp Tyr Gly Arg Leu Glu Val Gly Thr
        755                 760                 765
Arg Trp Leu Gly Asn Lys Leu Thr Leu Gly Gly Ala Met Arg Tyr Phe
    770                 775                 780
Gly Lys Ser Ile Arg Ala Thr Ala Glu Glu Arg Tyr Ile Asp Gly Thr
785                 790                 795                 800
Asn Gly Gly Asn Thr Ser Asn Phe Arg Gln Leu Gly Lys Arg Ser Ile
                805                 810                 815
Lys Gln Thr Glu Thr Leu Ala Arg Gln Pro Leu Ile Phe Asp Phe Tyr
            820                 825                 830
Ala Ala Tyr Glu Pro Lys Lys Asn Leu Ile Phe Arg Ala Glu Val Lys
        835                 840                 845
Asn Leu Phe Asp Arg Arg Tyr Ile Asp Pro Leu Asp Ala Gly Asn Asp
    850                 855                 860
Ala Ala Thr Gln Arg Tyr Tyr Ser Ser Phe Asp Pro Lys Asp Lys Asp
865                 870                 875                 880
Glu Asp Val Thr Cys Asn Ala Asp Lys Thr Leu Cys Asn Gly Lys Tyr
                885                 890                 895
Gly Gly Thr Ser Lys Ser Val Leu Thr Asn Phe Ala Arg Gly Arg Thr
            900                 905                 910
Phe Leu Ile Thr Met Ser Tyr Lys Phe
        915                 920

<210> 7
<211> 30
<212> DNA
<213> Bacteria

<400> 7
gctagctagc agatcttctt tcgggaagaa                                        30
```

<210> 8
<211> 34
<212> DNA
<213> Bacteria

<400> 8
ggtcgctcga gaaacttgta gctcatcgtt atca 34

<210> 9
<211> 1000
<212> DNA
<213> Bacteria

<400> 9
tatccgataa gtttccgtac cgaacagact agattcccgc ctgcgcggga atgacgattc 60
ataagtttcc cgaaattcca acataaccga aacctgacag taaccgtagc aactgaaccg 120
tcattcccac gaaagtggga atctagaaat gaaaagcaac aggcatttat cggaaataac 180
tgaaaccgaa cagactagat tcccgcctgc gcgggaatga cgattcataa gtttcccgaa 240
attccaacat aaccgaaacc tgacagtaac cgtagtaact gaaccgtcat tcccacgaaa 300
gtgggaatct agaaatgaaa agcaacaggc atttatcgga aataactgga accgaacaga 360
cctagattcc cgcctgcgcg ggaatgacgg ctgcagatgc ccgacggtct ttatagcaga 420
ttaacaaaaa tcaggacaag gcggcgagcc acagacagta caaacagtac ggaaccgatt 480
cacttggtgc ttcagcacct tagagaatcg ttctctttga gctaaggcga gacaacgccg 540
tactggtttt tgttaatccg ctatattccg ccatctctaa gatttacagc gatacacggg 600
taatttaagg aatggccgaa ccgtcattcc cgccactttt cgtcattccc accacttttc 660
gtcattcccg cgcaggcggg aatctagaat ctcggacttt cagataatct ttgaatattg 720
ctgttgttct aaggtttaga ttccccttgc gcgggaatga cgattcataa gtttcccgaa 780
attccaacat aaccgaaacc tgacagtaac cgtagcaaca gagaatcgtt ctctttgagc 840
taaggcgaaa caacgctgta ccggtttttg ttaatccact ataaatatcc aattgaaatc 900
ttcagacggt atatcgaatt tacacttttt aatgtttatg ccgcctaaaa aaatgctaat 960
atatttctta attgtctgac tgtttattgt tgaggaaaat 1000

BASB024 OUTER MEMBRANE PROTEIN OF NEISSERIA MENINGITIDIS

FIELD OF THE INVENTION

This invention relates to polynucleotides, (herein referred to as "BASB024 polynucleotide(s)"), polypeptides encoded by them (referred to herein as "BASB024" or "BASB024 polypeptide(s)"), recombinant materials and methods for their production. In another aspect, the invention relates to methods for using such polypeptides and polynucleotides, including vaccines against bacterial infections. In a further aspect, the invention relates to diagnostic assays for detecting infection of certain pathogens.

BACKGROUND OF THE INVENTION

Neisseria meningitidis (meningococcus) is a Gram-negative bacterium frequently isolated from the human upper respiratory tract. It occasionally causes invasive bacterial diseases such as bacteremia and meningitis. The incidence of meningococcal disease shows geographical seasonal and annual differences (Schwartz, B., Moore, P.S., Broome, C.V.; Clin. Microbiol. Rev. 2 (Supplement), S18-S24, 1989). Most disease in temperate countries is due to strains of serogroup B and varies in incidence from 1-10/100,000/year total population sometimes reaching higher values (Kaczmarski, E.B. (1997), Commun. Dis. Rep. Rev. 7: R55-9, 1995; Scholten, R.J.P.M., Bijlmer, H.A., Poolman, J.T. et al. Clin. Infect. Dis. 16: 237-246, 1993; Cruz, C., Pavez, G., Aguilar, E., et al. Epidemiol. Infect. 105: 119-126, 1990).

Epidemics dominated by serogroup A meningococci, mostly in central Africa, are encountered, sometimes reaching levels up to 1000/100.000/year (Schwartz, B., Moore, P.S., Broome, C.V. Clin. Microbiol. Rev. 2 (Supplement), S18-S24, 1989). Nearly all cases as a whole of meningococcal disease are caused by serogroup A, B, C, W-135 and Y meningococci and a tetravalent A, C, W-135, Y polysaccharide vaccine is available (Armand, J., Arminjon, F., Mynard, M.C., Lafaix, C., J. Biol. Stand. 10: 335-339, 1982).

The polysaccharide vaccines are currently being improved by way of chemical conjugating them to carrier proteins (Lieberman, J.M., Chiu, S.S., Wong, V.K., et al. JAMA 275 : 1499-1503, 1996).

A serogroup B vaccine is not available, since the B capsular polysaccharide was found to be nonimmunogenic, most likely because it shares structural similarity to host components (Wyle, F.A., Artenstein, M.S., Brandt, M.L. et al. J. Infect. Dis. 126: 514-522, 1972; Finne, J.M., Leinonen, M., Mäkelä, P.M. Lancet ii.: 355-357, 1983).

For many years efforts have been initiated and carried out to develop meningococcal outer membrane based vaccines (de Moraes, J.C., Perkins, B., Camargo, M.C. et al. Lancet 340: 1074-1078, 1992; Bjune, G., Hoiby, E.A. Gronnesby, J.K. et al. 338: 1093-1096, 1991). Such vaccines have demonstrated efficacies from 57% - 85% in older children (>4 years) and adolescents.

Many bacterial outer membrane components are present in these vaccines, such as PorA, PorB, Rmp, Opc, Opa, FrpB and the contribution of these components to the observed protection still needs futher definition. Other bacterial outer membrane components have been defined by using animal or human antibodies to be potentially relevant to the induction of protective immunity, such as TbpB and NspA (Martin, D., Cadieux, N., Hamel, J., Brodeux, B.R., J. Exp. Med. 185: 1173-1183, 1997; Lissolo, L., Maître-Wilmotte, C., Dumas, p. et al., Inf. Immun. 63: 884-890, 1995). The mechanisms of protective immunity will involve antibody mediated bactericidal activity and opsonophagocytosis.

A bacteremia animal model has been used to combine all antibody mediated mechanisms (Saukkonen, K., Leinonen, M., Abdillahi, H. Poolman, J. T. Vaccine 7: 325-328, 1989). It is generally accepted that the late complement component mediated bactericidal mechanism is

crucial for immunity against meningococcal disease (Ross, S.C., Rosenthal P.J., Berberic, H.M., Densen, P. J. Infect. Dis. 155: 1266-1275, 1987).

The frequency of *Neisseria meningitidis* infections has risen dramatically in the past few decades. This has been attributed to the emergence of multiply antibiotic resistant strains and an increasing population of people with weakened immune systems. It is no longer uncommon to isolate *Neisseria meningitidis* strains that are resistant to some or all of the standard antibiotics. This phenomenon has created an unmet medical need and demand for new anti-microbial agents, vaccines, drug screening methods, and diagnostic tests for this organism.

SUMMARY OF THE INVENTION

The present invention relates to BASB024, in particular BASB024 polypeptides and BASB024 polynucleotides, recombinant materials and methods for their production. In another aspect, the invention relates to methods for using such polypeptides and polynucleotides, including prevention and treatment of microbial diseases, amongst others. In a further aspect, the invention relates to diagnostic assays for detecting diseases associated with microbial infections and conditions associated with such infections, such as assays for detecting expression or activity of BASB024 polynucleotides or polypeptides.

Various changes and modifications within the spirit and scope of the disclosed invention will become readily apparent to those skilled in the art from reading the following descriptions and from reading the other parts of the present disclosure.

DESCRIPTION OF THE INVENTION

The invention relates to BASB024 polypeptides and polynucleotides as described in greater detail below. In particular, the invention relates to polypeptides and polynucleotides of BASB024 of *Neisseria meningitidis*, which is related by amino acid sequence homology to *Serratia marcescens* HasR heme acquisition outer membrane protein. The invention relates especially to BASB024 having the nucleotide and amino acid sequences set out in SEQ ID NO:1,3,5 and SEQ ID NO:2,4,6 respectively. It is understood that sequences recited in the Sequence Listing below as "DNA" represent an exemplification of one embodiment of the invention, since those of ordinary skill will recognize that such sequences can be usefully employed in polynucleotides in general, including ribopolynucleotides.

Polypeptides

In one aspect of the invention there are provided polypeptides of *Neisseria meningitidis* referred to herein as "BASB024" and "BASB024 polypeptides" as well as biologically, diagnostically, prophylactically, clinically or therapeutically useful variants thereof, and compositions comprising the same.

The present invention further provides for:

(a) an isolated polypeptide which comprises an amino acid sequence which has at least 85% identity, more preferably at least 90% identity, yet more preferably at least 95% identity, most preferably at least 97-99% or exact identity, to that of SEQ ID NO:2, 4, 6;

(b) a polypeptide encoded by an isolated polynucleotide comprising a polynucleotide sequence which has at least 85% identity, more preferably at least 90% identity, yet more preferably at least 95% identity, even more preferably at least 97-99% or exact identity to SEQ ID NO:1, 3, 5 over the entire length of SEQ ID NO:1, 3, 5 respectively; or

(c) a polypeptide encoded by an isolated polynucleotide comprising a polynucleotide sequence encoding a polypeptide which has at least 85% identity, more preferably at least 90% identity, yet more preferably at least 95% identity, even more preferably at least 97-99% or exact identity, to the amino acid sequence of SEQ ID NO:2, 4, 6;

The BASB024 polypeptides provided in SEQ ID NO:2,4,6 are the BASB024 polypeptides from *Neisseria meningitidis* strains ATCC13090 and H44/76.

The invention also provides an immunogenic fragment of a BASB024 polypeptide, that is, a contiguous portion of the BASB024 polypeptide which has the same or substantially the same immunogenic activity as the polypeptide comprising the amino acid sequence of SEQ ID NO:2,4,6. That is to say, the fragment (if necessary when coupled to a carrier) is capable of raising an immune response which recognises the BASB024 polypeptide. Such an immunogenic fragment may include, for example, the BASB024 polypeptide lacking an N-terminal leader sequence, and/or a transmembrane domain and/or a C-terminal anchor domain. In a preferred aspect the immunogenic fragment of BASB024 according to the invention comprises substantially all of the extracellular domain of a polypeptide which has at least 85% identity, more preferably at least 90% identity, yet more preferably at least 95% identity, most preferably at least 97-99% identity, to that of SEQ ID NO:2,4,6 over the entire length of SEQ ID NO:2

A fragment is a polypeptide having an amino acid sequence that is entirely the same as part but not all of any amino acid sequence of any polypeptide of the invention. As with BASB024 polypeptides, fragments may be "free-standing," or comprised within a larger polypeptide of which they form a part or region, most preferably as a single continuous region in a single larger polypeptide.

Preferred fragments include, for example, truncation polypeptides having a portion of an amino acid sequence of SEQ ID NO:2,4,6 or of variants thereof, such as a continuous series of residues that includes an amino- and/or carboxyl-terminal amino acid sequence. Degradation forms of the polypeptides of the invention produced by or in a host cell, are also preferred. Further preferred are fragments characterized by structural or functional attributes such as fragments that comprise alpha-helix and alpha-helix forming regions,

beta-sheet and beta-sheet-forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions.

Further preferred fragments include an isolated polypeptide comprising an amino acid sequence having at least 15, 20, 30, 40, 50 or 100 contiguous amino acids from the amino acid sequence of SEQ ID NO:2,4,6, or an isolated polypeptide comprising an amino acid sequence having at least 15, 20, 30, 40, 50 or 100 contiguous amino acids truncated or deleted from the amino acid sequence of SEQ ID NO:2,4,6.

Fragments of the polypeptides of the invention may be employed for producing the corresponding full-length polypeptide by peptide synthesis; therefore, these fragments may be employed as intermediates for producing the full-length polypeptides of the invention.

Particularly preferred are variants in which several, 5-10, 1-5, 1-3, 1-2 or 1 amino acids are substituted, deleted, or added in any combination.

The polypeptides, or immunogenic fragments, of the invention may be in the form of the "mature" protein or may be a part of a larger protein such as a precursor or a fusion protein. It is often advantageous to include an additional amino acid sequence which contains secretory or leader sequences, pro-sequences, sequences which aid in purification such as multiple histidine residues, or an additional sequence for stability during recombinant production. Furthermore, addition of exogenous polypeptide or lipid tail or polynucleotide sequences to increase the immunogenic potential of the final molecule is also considered.

In one aspect, the invention relates to genetically engineered soluble fusion proteins comprising a polypeptide of the present invention, or a fragment thereof, and various portions of the constant regions of heavy or light chains of immunoglobulins of various subclasses (IgG, IgM, IgA, IgE). Preferred as an immunoglobulin is the constant part of the heavy chain of human IgG, particularly IgG1, where fusion takes place at the hinge region. In a particular embodiment, the Fc part can be removed simply by incorporation of a cleavage sequence which can be cleaved with blood clotting factor Xa.

Furthermore, this invention relates to processes for the preparation of these fusion proteins by genetic engineering, and to the use thereof for drug screening, diagnosis and therapy. A further aspect of the invention also relates to polynucleotides encoding such fusion proteins. Examples of fusion protein technology can be found in International Patent Application Nos. WO94/29458 and WO94/22914.

The proteins may be chemically conjugated, or expressed as recombinant fusion proteins allowing increased levels to be produced in an expression system as compared to non-fused protein. The fusion partner may assist in providing T helper epitopes (immunological fusion partner), preferably T helper epitopes recognised by humans, or assist in expressing the protein (expression enhancer) at higher yields than the native recombinant protein. Preferably the fusion partner will be both an immunological fusion partner and expression enhancing partner.

Fusion partners include protein D from *Haemophilus influenzae* and the non-structural protein from influenzae virus, NS1 (hemagglutinin). Another fusion partner is the protein known as LytA. Preferably the C terminal portion of the molecule is used. LytA is derived from *Streptococcus pneumoniae* which synthesize an N-acetyl-L-alanine amidase, amidase LytA, (coded by the lytA gene {Gene, 43 (1986) page 265-272}) an autolysin that specifically degrades certain bonds in the peptidoglycan backbone. The C-terminal domain of the LytA protein is responsible for the affinity to

the choline or to some choline analogues such as DEAE. This property has been exploited for the development of *E.coli* C-LytA expressing plasmids useful for expression of fusion proteins. Purification of hybrid proteins containing the C-LytA fragment at its amino terminus has been described {Biotechnology: 10, (1992) page 795-798}. It is possible to use the repeat portion of the LytA molecule found in the C terminal end starting at residue 178, for example residues 188 - 305.

The present invention also includes variants of the aforementioned polypeptides, that is polypeptides that vary from the referents by conservative amino acid substitutions, whereby a residue is substituted by another with like characteristics. Typical such substitutions are among Ala, Val, Leu and Ile; among Ser and Thr; among the acidic residues Asp and Glu; among Asn and Gln; and among the basic residues Lys and Arg; or aromatic residues Phe and Tyr.

Polypeptides of the present invention can be prepared in any suitable manner. Such polypeptides include isolated naturally occurring polypeptides, recombinantly produced polypeptides, synthetically produced polypeptides, or polypeptides produced by a combination of these methods. Means for preparing such polypeptides are well understood in the art.

It is most preferred that a polypeptide of the invention is derived from *Neisseria meningitidis*, however, it may preferably be obtained from other organisms of the same taxonomic genus. A polypeptide of the invention may also be obtained, for example, from organisms of the same taxonomic family or order.

Polynucleotides

It is an object of the invention to provide polynucleotides that encode BASB024 polypeptides, particularly polynucleotides that encode the polypeptide herein designated BASB024.

In a particularly preferred embodiment of the invention the polynucleotide comprises a region encoding BASB024 polypeptides comprising a sequence set out in SEQ ID NO:1,3,5 which includes a full length gene, or a variant thereof.

The BASB024 polynucleotides provided in SEQ ID NO:1,3,5 are the BASB024 polynucleotides from *Neisseria meningitidis* strains ATCC13090 and H44/76.

As a further aspect of the invention there are provided isolated nucleic acid molecules encoding and/or expressing BASB024 polypeptides and polynucleotides, particularly *Neisseria meningitidis* BASB024 polypeptides and polynucleotides, including, for example, unprocessed RNAs, ribozyme RNAs, mRNAs, cDNAs, genomic DNAs, B- and Z-DNAs. Further embodiments of the invention include biologically, diagnostically, prophylactically, clinically or therapeutically useful polynucleotides and polypeptides, and variants thereof, and compositions comprising the same.

Another aspect of the invention relates to isolated polynucleotides, including at least one full length gene, that encodes a BASB024 polypeptide having a deduced amino acid sequence of SEQ ID NO:2,4,6 and polynucleotides closely related thereto and variants thereof.

In another particularly preferred embodiment of the invention there is a BASB024 polypeptide from *Neisseria meningitidis* comprising or consisting of an amino acid sequence of SEQ ID NO:2,4,6 or a variant thereof.

Using the information provided herein, such as a polynucleotide sequence set out in SEQ ID NO:1, 3, 5 a polynucleotide of the invention encoding BASB024 polypeptide may be obtained using standard cloning and screening methods, such as those for cloning and sequencing chromosomal DNA fragments from bacteria using *Neisseria meningitidis* cells as starting material, followed by obtaining a full length clone. For example, to obtain a

polynucleotide sequence of the invention, such as a polynucleotide sequence given in SEQ ID NO:1,3,5, typically a library of clones of chromosomal DNA of *Neisseria meningitidis* in *E.coli* or some other suitable host is probed with a radiolabeled oligonucleotide, preferably a 17-mer or longer, derived from a partial sequence. Clones carrying DNA identical to that of the probe can then be distinguished using stringent hybridization conditions. By sequencing the individual clones thus identified by hybridization with sequencing primers designed from the original polypeptide or polynucleotide sequence it is then possible to extend the polynucleotide sequence in both directions to determine a full length gene sequence. Conveniently, such sequencing is performed, for example, using denatured double stranded DNA prepared from a plasmid clone. Suitable techniques are described by Maniatis, T., Fritsch, E.F. and Sambrook et al., *MOLECULAR CLONING, A LABORATORY MANUAL*, 2nd Ed.; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1989). (see in particular Screening By Hybridization 1.90 and Sequencing Denatured Double-Stranded DNA Templates 13.70). Direct genomic DNA sequencing may also be performed to obtain a full length gene sequence. Illustrative of the invention, each polynucleotide set out in SEQ ID NO:1,3,5 was discovered in a DNA library derived from *Neisseria meningitidis*.

Moreover, each DNA sequence set out in SEQ ID NO:1,3,5 contains an open reading frame encoding a protein having about the number of amino acid residues set forth in SEQ ID NO:2, 4, 6 with a deduced molecular weight that can be calculated using amino acid residue molecular weight values well known to those skilled in the art.

The polynucleotide of SEQ ID NO:1, between the start codon at nucleotide number 1 and the stop codon which begins at nucleotide number 2767 of SEQ ID NO:1, encodes the polypeptide of SEQ ID NO:2.

The polynucleotide of SEQ ID NO:3, between the start codon at nucleotide number 1 and the stop codon which begins at nucleotide number 2767 of SEQ ID NO:3, encodes the polypeptide of SEQ ID NO:4.

The polynucleotide of SEQ ID NO:5, between the start codon at nucleotide number 1 and the stop codon which begins at nucleotide number 2764 of SEQ ID NO:5, encodes the polypeptide of SEQ ID NO:6.

In a further aspect, the present invention provides for an isolated polynucleotide comprising or consisting of:

(a) a polynucleotide sequence which has at least 85% identity, more preferably at least 90% identity, yet more preferably at least 95% identity, even more preferably at least 97-99% or exact identity to SEQ ID NO:1,3,5 over the entire length of SEQ ID NO:1,3,5 respectively; or

(b) a polynucleotide sequence encoding a polypeptide which has at least 85% identity, more preferably at least 90% identity, yet more preferably at least 95% identity, even more preferably at least 97-99% or 100% exact, to the amino acid sequence of SEQ ID NO:2, 4, 6 over the entire length of SEQ ID NO:2, 4, 6 respectively.

A polynucleotide encoding a polypeptide of the present invention, including homologs and orthologs from species other than *Neisseria meningitidis*, may be obtained by a process which comprises the steps of screening an appropriate library under stringent hybridization conditions (for example, using a temperature in the range of 45 – 65°C and an SDS concentration from 0.1 – 1%) with a labeled or detectable probe consisting of or comprising the sequence of SEQ ID NO: 1, 3, 5 or a fragment thereof; and isolating a full-length gene and/or genomic clones containing said polynucleotide sequence.

The invention provides a polynucleotide sequence identical over its entire length to a coding sequence (open reading frame) in SEQ ID NO: 1, 3, 5. Also provided by the invention is a coding sequence for a mature polypeptide or a fragment thereof, by itself as well as a coding sequence for a mature polypeptide or a fragment in reading frame with another coding sequence, such as a sequence encoding a leader or secretory sequence, a pre-, or pro- or prepro-protein sequence. The polynucleotide of the invention may also contain at least one

non-coding sequence, including for example, but not limited to at least one non-coding 5' and 3' sequence, such as the transcribed but non-translated sequences, termination signals (such as rho-dependent and rho-independent termination signals), ribosome binding sites, Kozak sequences, sequences that stabilize mRNA, introns, and polyadenylation signals. The polynucleotide sequence may also comprise additional coding sequence encoding additional amino acids. For example, a marker sequence that facilitates purification of the fused polypeptide can be encoded. In certain embodiments of the invention, the marker sequence is a hexa-histidine peptide, as provided in the pQE vector (Qiagen, Inc.) and described in Gentz *et al.*, *Proc. Natl. Acad. Sci., USA 86:* 821-824 (1989), or an HA peptide tag (Wilson *et al.*, *Cell 37:* 767 (1984), both of which may be useful in purifying polypeptide sequence fused to them. Polynucleotides of the invention also include, but are not limited to, polynucleotides comprising a structural gene and its naturally associated sequences that control gene expression.

The nucleotide sequence encoding BASB024 polypeptide of SEQ ID NO:2, 4, 6 may be identical to the polypeptide encoding sequence contained in nucleotides 1 to 2766 of SEQ ID NO:1, or the polypeptide encoding sequence contained in nucleotides 1 to 2766 of SEQ ID NO:3, or the polypeptide encoding sequence contained in nucleotides 1 to 2763 of SEQ ID NO:5, respectively. Alternatively it may be a sequence, which as a result of the redundancy (degeneracy) of the genetic code, also encodes the polypeptide of SEQ ID NO:2, 4, 6.

The term "polynucleotide encoding a polypeptide" as used herein encompasses polynucleotides that include a sequence encoding a polypeptide of the invention, particularly a bacterial polypeptide and more particularly a polypeptide of the *Neisseria meningitidis* BASB024 having an amino acid sequence set out in SEQ ID NO:2, 4, 6. The term also encompasses polynucleotides that include a single continuous region or discontinuous regions encoding the polypeptide (for example, polynucleotides interrupted by integrated phage, an integrated insertion sequence, an integrated vector sequence, an

integrated transposon sequence, or due to RNA editing or genomic DNA reorganization) together with additional regions, that also may contain coding and/or non-coding sequences.

The invention further relates to variants of the polynucleotides described herein that encode variants of a polypeptide having a deduced amino acid sequence of SEQ ID NO:2, 4, 6. Fragments of polynucleotides of the invention may be used, for example, to synthesize full-length polynucleotides of the invention.

Further particularly preferred embodiments are polynucleotides encoding BASB024 variants, that have the amino acid sequence of BASB024 polypeptide of SEQ ID NO:2, 4, 6 in which several, a few, 5 to 10, 1 to 5, 1 to 3, 2, 1 or no amino acid residues are substituted, modified, deleted and/or added, in any combination. Especially preferred among these are silent substitutions, additions and deletions, that do not alter the properties and activities of BASB024 polypeptide.

Further preferred embodiments of the invention are polynucleotides that are at least 85% identical over their entire length to a polynucleotide encoding BASB024 polypeptide having an amino acid sequence set out in SEQ ID NO:2, 4, 6, and polynucleotides that are complementary to such polynucleotides. In this regard, polynucleotides at least 90% identical over their entire length to the same are particularly preferred, and among these particularly preferred polynucleotides, those with at least 95% are especially preferred. Furthermore, those with at least 97% are highly preferred among those with at least 95%, and among these those with at least 98% and at least 99% are particularly highly preferred, with at least 99% being the more preferred.

Preferred embodiments are polynucleotides encoding polypeptides that retain substantially the same biological function or activity as the mature polypeptide encoded by a DNA of SEQ ID NO:1, 3, 5.

In accordance with certain preferred embodiments of this invention there are provided polynucleotides that hybridize, particularly under stringent conditions, to BASB024 polynucleotide sequences, such as those polynucleotides in SEQ ID NO:1, 3, 5.

The invention further relates to polynucleotides that hybridize to the polynucleotide sequences provided herein. In this regard, the invention especially relates to polynucleotides that hybridize under stringent conditions to the polynucleotides described herein. As herein used, the terms "stringent conditions" and "stringent hybridization conditions" mean hybridization occurring only if there is at least 95% and preferably at least 97% identity between the sequences. A specific example of stringent hybridization conditions is overnight incubation at 42°C in a solution comprising: 50% formamide, 5x SSC (150mM NaCl, 15mM trisodium citrate), 50 mM sodium phosphate (pH7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 micrograms/ml of denatured, sheared salmon sperm DNA, followed by washing the hybridization support in 0.1x SSC at about 65°C. Hybridization and wash conditions are well known and exemplified in Sambrook, *et al.*, Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, N.Y., (1989), particularly Chapter 11 therein. Solution hybridization may also be used with the polynucleotide sequences provided by the invention.

The invention also provides a polynucleotide consisting of or comprising a polynucleotide sequence obtained by screening an appropriate library containing the complete gene for a polynucleotide sequence set forth in SEQ ID NO:1, 3, 5 under stringent hybridization conditions with a probe having the sequence of said polynucleotide sequence set forth in SEQ ID NO:1, 3, 5 or a fragment thereof; and isolating said polynucleotide sequence. Fragments useful for obtaining such a polynucleotide include, for example, probes and primers fully described elsewhere herein.

As discussed elsewhere herein regarding polynucleotide assays of the invention, for instance, the polynucleotides of the invention, may be used as a hybridization probe for

RNA, cDNA and genomic DNA to isolate full-length cDNAs and genomic clones encoding BASB024 and to isolate cDNA and genomic clones of other genes that have a high identity, particularly high sequence identity, to the BASB024 gene. Such probes generally will comprise at least 15 nucleotide residues or base pairs. Preferably, such probes will have at least 30 nucleotide residues or base pairs and may have at least 50 nucleotide residues or base pairs. Particularly preferred probes will have at least 20 nucleotide residues or base pairs and will have less than 30 nucleotide residues or base pairs.

A coding region of a BASB024 gene may be isolated by screening using a DNA sequence provided in SEQ ID NO:1, 3, 5 to synthesize an oligonucleotide probe. A labeled oligonucleotide having a sequence complementary to that of a gene of the invention is then used to screen a library of cDNA, genomic DNA or mRNA to determine which members of the library the probe hybridizes to.

There are several methods available and well known to those skilled in the art to obtain full-length DNAs, or extend short DNAs, for example those based on the method of Rapid Amplification of cDNA ends (RACE) (see, for example, Frohman, *et al.*, *PNAS USA 85:* 8998-9002, 1988). Recent modifications of the technique, exemplified by the Marathon™ technology (Clontech Laboratories Inc.) for example, have significantly simplified the search for longer cDNAs. In the Marathon™ technology, cDNAs have been prepared from mRNA extracted from a chosen tissue and an 'adaptor' sequence ligated onto each end. Nucleic acid amplification (PCR) is then carried out to amplify the "missing" 5' end of the DNA using a combination of gene specific and adaptor specific oligonucleotide primers. The PCR reaction is then repeated using "nested" primers, that is, primers designed to anneal within the amplified product (typically an adaptor specific primer that anneals further 3' in the adaptor sequence and a gene specific primer that anneals further 5' in the selected gene sequence). The products of this reaction can then be analyzed by DNA sequencing and a full-length DNA constructed either by joining the product directly

to the existing DNA to give a complete sequence, or carrying out a separate full-length PCR using the new sequence information for the design of the 5' primer.

The polynucleotides and polypeptides of the invention may be employed, for example, as research reagents and materials for discovery of treatments of and diagnostics for diseases, particularly human diseases, as further discussed herein relating to polynucleotide assays.

The polynucleotides of the invention that are oligonucleotides derived from a sequence of SEQ ID NOS:1 – 6 may be used in the processes herein as described, but preferably for PCR, to determine whether or not the polynucleotides identified herein in whole or in part are transcribed in bacteria in infected tissue. It is recognized that such sequences will also have utility in diagnosis of the stage of infection and type of infection the pathogen has attained.

The invention also provides polynucleotides that encode a polypeptide that is the mature protein plus additional amino or carboxyl-terminal amino acids, or amino acids interior to the mature polypeptide (when the mature form has more than one polypeptide chain, for instance). Such sequences may play a role in processing of a protein from precursor to a mature form, may allow protein transport, may lengthen or shorten protein half-life or may facilitate manipulation of a protein for assay or production, among other things. As generally is the case *in vivo*, the additional amino acids may be processed away from the mature protein by cellular enzymes.

For each and every polynucleotide of the invention there is provided a polynucleotide complementary to it. It is preferred that these complementary polynucleotides are fully complementary to each polynucleotide with which they are complementary.

A precursor protein, having a mature form of the polypeptide fused to one or more prosequences may be an inactive form of the polypeptide. When prosequences are removed

such inactive precursors generally are activated. Some or all of the prosequences may be removed before activation. Generally, such precursors are called proproteins.

In addition to the standard A, G, C, T/U representations for nucleotides, the term "N" may also be used in describing certain polynucleotides of the invention. "N" means that any of the four DNA or RNA nucleotides may appear at such a designated position in the DNA or RNA sequence, except it is preferred that N is not a nucleic acid that when taken in combination with adjacent nucleotide positions, when read in the correct reading frame, would have the effect of generating a premature termination codon in such reading frame.

In sum, a polynucleotide of the invention may encode a mature protein, a mature protein plus a leader sequence (which may be referred to as a preprotein), a precursor of a mature protein having one or more prosequences that are not the leader sequences of a preprotein, or a preproprotein, which is a precursor to a proprotein, having a leader sequence and one or more prosequences, which generally are removed during processing steps that produce active and mature forms of the polypeptide.

In accordance with an aspect of the invention, there is provided the use of a polynucleotide of the invention for therapeutic or prophylactic purposes, in particular genetic immunization.

The use of a polynucleotide of the invention in genetic immunization will preferably employ a suitable delivery method such as direct injection of plasmid DNA into muscles (Wolff *et al., Hum Mol Genet* (1992) 1: 363, Manthorpe *et al., Hum. Gene Ther.* (1983) 4: 419), delivery of DNA complexed with specific protein carriers (Wu *et al., J Biol Chem.* (1989) 264: 16985), coprecipitation of DNA with calcium phosphate (Benvenisty & Reshef, *PNAS USA*, (1986) 83: 9551), encapsulation of DNA in various forms of liposomes (Kaneda *et al., Science* (1989) 243: 375), particle bombardment (Tang *et al.,*

Nature (1992) 356:152, Eisenbraun *et al., DNA Cell Biol* (1993) 12: 791) and *in vivo* infection using cloned retroviral vectors (Seeger *et al.*, *PNAS USA* (1984) 81: 5849).

Vectors, Host Cells, Expression Systems

The invention also relates to vectors that comprise a polynucleotide or polynucleotides of the invention, host cells that are genetically engineered with vectors of the invention and the production of polypeptides of the invention by recombinant techniques. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the invention.

Recombinant polypeptides of the present invention may be prepared by processes well known in those skilled in the art from genetically engineered host cells comprising expression systems. Accordingly, in a further aspect, the present invention relates to expression systems that comprise a polynucleotide or polynucleotides of the present invention, to host cells which are genetically engineered with such expression systems, and to the production of polypeptides of the invention by recombinant techniques.

For recombinant production of the polypeptides of the invention, host cells can be genetically engineered to incorporate expression systems or portions thereof or polynucleotides of the invention. Introduction of a polynucleotide into the host cell can be effected by methods described in many standard laboratory manuals, such as Davis, *et al.*, *BASIC METHODS IN MOLECULAR BIOLOGY*, (1986) and Sambrook, *et al.*, *MOLECULAR CLONING: A LABORATORY MANUAL*, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989), such as, calcium phosphate transfection, DEAE-dextran mediated transfection, transvection, microinjection, cationic lipid-mediated transfection, electroporation, transduction, scrape loading, ballistic introduction and infection.

Representative examples of appropriate hosts include bacterial cells, such as cells of streptococci, staphylococci, enterococci, *E. coli*, streptomyces, cyanobacteria, *Bacillus subtilis*, *Moraxella catarrhalis, Haemophilus influenzae* and *Neisseria meningitidis*; fungal cells, such as cells of a yeast, *Kluveromyces*, *Saccharomyces*, a basidiomycete, *Candida albicans* and *Aspergillus*; insect cells such as cells of *Drosophila* S2 and *Spodoptera* Sf9; animal cells such as CHO, COS, HeLa, C127, 3T3, BHK, 293, CV-1 and Bowes melanoma cells; and plant cells, such as cells of a gymnosperm or angiosperm.

A great variety of expression systems can be used to produce the polypeptides of the invention. Such vectors include, among others, chromosomal-, episomal- and virus-derived vectors, for example, vectors derived from bacterial plasmids, from bacteriophage, from transposons, from yeast episomes, from insertion elements, from yeast chromosomal elements, from viruses such as baculoviruses, papova viruses, such as SV40, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses, picornaviruses, retroviruses, and alphaviruses and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids. The expression system constructs may contain control regions that regulate as well as engender expression. Generally, any system or vector suitable to maintain, propagate or express polynucleotides and/or to express a polypeptide in a host may be used for expression in this regard. The appropriate DNA sequence may be inserted into the expression system by any of a variety of well-known and routine techniques, such as, for example, those set forth in Sambrook *et al.*, *MOLECULAR CLONING, A LABORATORY MANUAL*, (*supra*).

In recombinant expression systems in eukaryotes, for secretion of a translated protein into the lumen of the endoplasmic reticulum, into the periplasmic space or into the extracellular environment, appropriate secretion signals may be incorporated into the expressed polypeptide. These signals may be endogenous to the polypeptide or they may be heterologous signals.

Polypeptides of the present invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, ion metal affinity chromatography (IMAC) is employed for purification. Well known techniques for refolding proteins may be employed to regenerate active conformation when the polypeptide is denatured during intracellular synthesis, isolation and or purification.

The expression system may also be a recombinant live microorganism, such as a virus or bacterium. The gene of interest can be inserted into the genome of a live recombinant virus or bacterium. Inoculation and *in vivo* infection with this live vector will lead to *in vivo* expression of the antigen and induction of immune responses. Viruses and bacteria used for this purpose are for instance: poxviruses (e.g; vaccinia, fowlpox, canarypox), alphaviruses (Sindbis virus, Semliki Forest Virus, Venezuelian Equine Encephalitis Virus), adenoviruses, adeno-associated virus, picornaviruses (poliovirus, rhinovirus), herpesviruses (varicella zoster virus, etc), Listeria, Salmonella , Shigella, Neisseria, BCG. These viruses and bacteria can be virulent, or attenuated in various ways in order to obtain live vaccines. Such live vaccines also form part of the invention.

Diagnostic, Prognostic, Serotyping and Mutation Assays

This invention is also related to the use of BASB024 polynucleotides and polypeptides of the invention for use as diagnostic reagents. Detection of BASB024 polynucleotides and/or polypeptides in a eukaryote, particularly a mammal, and especially a human, will provide a diagnostic method for diagnosis of disease, staging of disease or response of an infectious organism to drugs. Eukaryotes, particularly mammals, and especially humans, particularly those infected or suspected to be infected with an organism comprising the BASB024 gene or protein, may be detected at the nucleic acid or amino acid level by a variety of well known techniques as well as by methods provided herein.

Polypeptides and polynucleotides for prognosis, diagnosis or other analysis may be obtained from a putatively infected and/or infected individual's bodily materials. Polynucleotides from any of these sources, particularly DNA or RNA, may be used directly for detection or may be amplified enzymatically by using PCR or any other amplification technique prior to analysis. RNA, particularly mRNA, cDNA and genomic DNA may also be used in the same ways. Using amplification, characterization of the species and strain of infectious or resident organism present in an individual, may be made by an analysis of the genotype of a selected polynucleotide of the organism. Deletions and insertions can be detected by a change in size of the amplified product in comparison to a genotype of a reference sequence selected from a related organism, preferably a different species of the same genus or a different strain of the same species. Point mutations can be identified by hybridizing amplified DNA to labeled BASB024 polynucleotide sequences. Perfectly or significantly matched sequences can be distinguished from imperfectly or more significantly mismatched duplexes by DNase or RNase digestion, for DNA or RNA respectively, or by detecting differences in melting temperatures or renaturation kinetics. Polynucleotide sequence differences may also be detected by alterations in the electrophoretic mobility of polynucleotide fragments in gels as compared to a reference sequence. This may be carried out with or without denaturing agents. Polynucleotide differences may also be detected by direct DNA or RNA sequencing. See, for example, Myers *et al.*, *Science, 230:* 1242 (1985). Sequence changes at specific locations also may be revealed by nuclease protection assays, such as RNase, V1 and S1 protection assay or a chemical cleavage method. See, for example, Cotton *et al.*, *Proc. Natl. Acad. Sci., USA, 85:* 4397-4401 (1985).

In another embodiment, an array of oligonucleotides probes comprising BASB024 nucleotide sequence or fragments thereof can be constructed to conduct efficient screening of, for example, genetic mutations, serotype, taxonomic classification or identification. Array technology methods are well known and have general applicability and can be used to

address a variety of questions in molecular genetics including gene expression, genetic linkage, and genetic variability (see, for example, Chee *et al.*, *Science, 274: 610* (1996)).

Thus in another aspect, the present invention relates to a diagnostic kit which comprises:

(a) a polynucleotide of the present invention, preferably the nucleotide sequence of SEQ ID NO:1, 3, 5, or a fragment thereof;

(b) a nucleotide sequence complementary to that of (a);

(c) a polypeptide of the present invention, preferably the polypeptide of SEQ ID NO:2, 4, 6 or a fragment thereof; or

(d) an antibody to a polypeptide of the present invention, preferably to the polypeptide of SEQ ID NO:2, 4, 6.

It will be appreciated that in any such kit, (a), (b), (c) or (d) may comprise a substantial component. Such a kit will be of use in diagnosing a disease or susceptibility to a disease, among others.

This invention also relates to the use of polynucleotides of the present invention as diagnostic reagents. Detection of a mutated form of a polynucleotide of the invention, preferable, SEQ ID NO:1, 3, 5, which is associated with a disease or pathogenicity will provide a diagnostic tool that can add to, or define, a diagnosis of a disease, a prognosis of a course of disease, a determination of a stage of disease, or a susceptibility to a disease, which results from under-expression, over-expression or altered expression of the polynucleotide. Organisms, particularly infectious organisms, carrying mutations in such polynucleotide may be detected at the polynucleotide level by a variety of techniques, such as those described elsewhere herein.

Cells from an organism carrying mutations or polymorphisms (allelic variations) in a polynucleotide and/or polypeptide of the invention may also be detected at the polynucleotide or polypeptide level by a variety of techniques, to allow for serotyping, for

example. For example, RT-PCR can be used to detect mutations in the RNA. It is particularly preferred to use RT-PCR in conjunction with automated detection systems, such as, for example, GeneScan. RNA, cDNA or genomic DNA may also be used for the same purpose, PCR. As an example, PCR primers complementary to a polynucleotide encoding BASB024 polypeptide can be used to identify and analyze mutations.

The invention further provides primers with 1, 2, 3 or 4 nucleotides removed from the 5' and/or the 3' end. These primers may be used for, among other things, amplifying BASB024 DNA and/or RNA isolated from a sample derived from an individual, such as a bodily material. The primers may be used to amplify a polynucleotide isolated from an infected individual, such that the polynucleotide may then be subject to various techniques for elucidation of the polynucleotide sequence. In this way, mutations in the polynucleotide sequence may be detected and used to diagnose and/or prognose the infection or its stage or course, or to serotype and/or classify the infectious agent.

The invention further provides a process for diagnosing disease, preferably bacterial infections, more preferably infections caused by *Neisseria meningitidis*, comprising determining from a sample derived from an individual, such as a bodily material, an increased level of expression of polynucleotide having a sequence of SEQ ID NO:1, 3, 5. Increased or decreased expression of a BASB024 polynucleotide can be measured using any on of the methods well known in the art for the quantitation of polynucleotides, such as, for example, amplification, PCR, RT-PCR, RNase protection, Northern blotting, spectrometry and other hybridization methods.

In addition, a diagnostic assay in accordance with the invention for detecting over-expression of BASB024 polypeptide compared to normal control tissue samples may be used to detect the presence of an infection, for example. Assay techniques that can be used to determine levels of a BASB024 polypeptide, in a sample derived from a host, such as a bodily material, are well-known to those of skill in the art. Such assay methods include

radioimmunoassays, competitive-binding assays, Western Blot analysis, antibody sandwich assays, antibody detection and ELISA assays.

The polynucleotides of the invention may be used as components of polynucleotide arrays, preferably high density arrays or grids. These high density arrays are particularly useful for diagnostic and prognostic purposes. For example, a set of spots each comprising a different gene, and further comprising a polynucleotide or polynucleotides of the invention, may be used for probing, such as using hybridization or nucleic acid amplification, using a probe obtained or derived from a bodily sample, to determine the presence of a particular polynucleotide sequence or related sequence in an individual. Such a presence may indicate the presence of a pathogen, particularly *Neisseria meningitidis*, and may be useful in diagnosing and/or prognosing disease or a course of disease. A grid comprising a number of variants of the polynucleotide sequence of SEQ ID NO:1, 3, 5 are preferred. Also preferred is a grid comprising a number of variants of a polynucleotide sequence encoding the polypeptide sequence of SEQ ID NO:2, 4, 6.

Antibodies

The polypeptides and polynucleotides of the invention or variants thereof, or cells expressing the same can be used as immunogens to produce antibodies immunospecific for such polypeptides or polynucleotides respectively.

In certain preferred embodiments of the invention there are provided antibodies against BASB024 polypeptides or polynucleotides.

Antibodies generated against the polypeptides or polynucleotides of the invention can be obtained by administering the polypeptides and/or polynucleotides of the invention, or epitope-bearing fragments of either or both, analogues of either or both, or cells expressing either or both, to an animal, preferably a nonhuman, using routine protocols. For

preparation of monoclonal antibodies, any technique known in the art that provides antibodies produced by continuous cell line cultures can be used. Examples include various techniques, such as those in Kohler, G. and Milstein, C., *Nature 256:* 495-497 (1975); Kozbor *et al.*, *Immunology Today 4:* 72 (1983); Cole *et al.*, pg. 77-96 in *MONOCLONAL ANTIBODIES AND CANCER THERAPY*, Alan R. Liss, Inc. (1985).

Techniques for the production of single chain antibodies (U.S. Patent No. 4,946,778) can be adapted to produce single chain antibodies to polypeptides or polynucleotides of this invention. Also, transgenic mice, or other organisms or animals, such as other mammals, may be used to express humanized antibodies immunospecific to the polypeptides or polynucleotides of the invention.

Alternatively, phage display technology may be utilized to select antibody genes with binding activities towards a polypeptide of the invention either from repertoires of PCR amplified v-genes of lymphocytes from humans screened for possessing anti-BASB024 or from naive libraries (McCafferty, *et al.*, (1990), Nature 348, 552-554; Marks, *et al.*, (1992) *Biotechnology 10*, 779-783). The affinity of these antibodies can also be improved by, for example, chain shuffling (Clackson *et al.*, (1991) *Nature 352:* 628).

The above-described antibodies may be employed to isolate or to identify clones expressing the polypeptides or polynucleotides of the invention to purify the polypeptides or polynucleotides by, for example, affinity chromatography.

Thus, among others, antibodies against BASB024-polypeptide or BASB024-polynucleotide may be employed to treat infections, particularly bacterial infections.

Polypeptide variants include antigenically, epitopically or immunologically equivalent variants form a particular aspect of this invention.

Preferably, the antibody or variant thereof is modified to make it less immunogenic in the individual. For example, if the individual is human the antibody may most preferably be "humanized," where the complimentarity determining region or regions of the hybridoma-derived antibody has been transplanted into a human monoclonal antibody, for example as described in Jones *et al.* (1986), *Nature* 321, 522-525 or Tempest *et al.*, (1991) *Biotechnology* 9, 266-273.

Antagonists and Agonists - Assays and Molecules

Polypeptides and polynucleotides of the invention may also be used to assess the binding of small molecule substrates and ligands in, for example, cells, cell-free preparations, chemical libraries, and natural product mixtures. These substrates and ligands may be natural substrates and ligands or may be structural or functional mimetics. See, *e.g.*, Coligan *et al.*, *Current Protocols in Immunology 1(2):* Chapter 5 (1991).

The screening methods may simply measure the binding of a candidate compound to the polypeptide or polynucleotide, or to cells or membranes bearing the polypeptide or polynucleotide, or a fusion protein of the polypeptide by means of a label directly or indirectly associated with the candidate compound. Alternatively, the screening method may involve competition with a labeled competitor. Further, these screening methods may test whether the candidate compound results in a signal generated by activation or inhibition of the polypeptide or polynucleotide, using detection systems appropriate to the cells comprising the polypeptide or polynucleotide. Inhibitors of activation are generally assayed in the presence of a known agonist and the effect on activation by the agonist by the presence of the candidate compound is observed. Constitutively active polypeptide and/or constitutively expressed polypeptides and polynucleotides may be employed in screening methods for inverse agonists or inhibitors, in the absence of an agonist or inhibitor, by testing whether the candidate compound results in inhibition of activation of the polypeptide or polynucleotide, as the case may be. Further, the screening methods

may simply comprise the steps of mixing a candidate compound with a solution containing a polypeptide or polynucleotide of the present invention, to form a mixture, measuring BASB024 polypeptide and/or polynucleotide activity in the mixture, and comparing the BASB024 polypeptide and/or polynucleotide activity of the mixture to a standard. Fusion proteins, such as those made from Fc portion and BASB024 polypeptide, as hereinbefore described, can also be used for high-throughput screening assays to identify antagonists of the polypeptide of the present invention, as well as of phylogenetically and and/or functionally related polypeptides (see D. Bennett *et al.*, J Mol Recognition, 8:52-58 (1995); and K. Johanson *et al.*, J Biol Chem, 270(16):9459-9471 (1995)).

The polynucleotides, polypeptides and antibodies that bind to and/or interact with a polypeptide of the present invention may also be used to configure screening methods for detecting the effect of added compounds on the production of mRNA and/or polypeptide in cells. For example, an ELISA assay may be constructed for measuring secreted or cell associated levels of polypeptide using monoclonal and polyclonal antibodies by standard methods known in the art. This can be used to discover agents which may inhibit or enhance the production of polypeptide (also called antagonist or agonist, respectively) from suitably manipulated cells or tissues.

The invention also provides a method of screening compounds to identify those which enhance (agonist) or block (antagonist) the action of BASB024 polypeptides or polynucleotides, particularly those compounds that are bacteristatic and/or bactericidal. The method of screening may involve high-throughput techniques. For example, to screen for agonists or antagonists, a synthetic reaction mix, a cellular compartment, such as a membrane, cell envelope or cell wall, or a preparation of any thereof, comprising BASB024 polypeptide and a labeled substrate or ligand of such polypeptide is incubated in the absence or the presence of a candidate molecule that may be a BASB024 agonist or antagonist. The ability of the candidate molecule to agonize or antagonize the BASB024 polypeptide is

reflected in decreased binding of the labeled ligand or decreased production of product from such substrate. Molecules that bind gratuitously, *i.e.*, without inducing the effects of BASB024 polypeptide are most likely to be good antagonists. Molecules that bind well and, as the case may be, increase the rate of product production from substrate, increase signal transduction, or increase chemical channel activity are agonists. Detection of the rate or level of, as the case may be, production of product from substrate, signal transduction, or chemical channel activity may be enhanced by using a reporter system. Reporter systems that may be useful in this regard include but are not limited to colorimetric, labeled substrate converted into product, a reporter gene that is responsive to changes in BASB024 polynucleotide or polypeptide activity, and binding assays known in the art.

Another example of an assay for BASB024 agonists is a competitive assay that combines BASB024 and a potential agonist with BASB024-binding molecules, recombinant BASB024 binding molecules, natural substrates or ligands, or substrate or ligand mimetics, under appropriate conditions for a competitive inhibition assay. BASB024 can be labeled, such as by radioactivity or a colorimetric compound, such that the number of BASB024 molecules bound to a binding molecule or converted to product can be determined accurately to assess the effectiveness of the potential antagonist.

Potential antagonists include, among others, small organic molecules, peptides, polypeptides and antibodies that bind to a polynucleotide and/or polypeptide of the invention and thereby inhibit or extinguish its activity or expression. Potential antagonists also may be small organic molecules, a peptide, a polypeptide such as a closely related protein or antibody that binds the same sites on a binding molecule, such as a binding molecule, without inducing BASB024-induced activities, thereby preventing the action or expression of BASB024 polypeptides and/or polynucleotides by excluding BASB024 polypeptides and/or polynucleotides from binding.

Potential antagonists include a small molecule that binds to and occupies the binding site of the polypeptide thereby preventing binding to cellular binding molecules, such that normal biological activity is prevented. Examples of small molecules include but are not limited to small organic molecules, peptides or peptide-like molecules. Other potential antagonists include antisense molecules (see Okano, *J. Neurochem. 56:* 560 (1991); *OLIGODEOXYNUCLEOTIDES AS ANTISENSE INHIBITORS OF GENE EXPRESSION*, CRC Press, Boca Raton, FL (1988), for a description of these molecules). Preferred potential antagonists include compounds related to and variants of BASB024.

In a further aspect, the present invention relates to genetically engineered soluble fusion proteins comprising a polypeptide of the present invention, or a fragment thereof, and various portions of the constant regions of heavy or light chains of immunoglobulins of various subclasses (IgG, IgM, IgA, IgE). Preferred as an immunoglobulin is the constant part of the heavy chain of human IgG, particularly IgG1, where fusion takes place at the hinge region. In a particular embodiment, the Fc part can be removed simply by incorporation of a cleavage sequence which can be cleaved with blood clotting factor Xa. Furthermore, this invention relates to processes for the preparation of these fusion proteins by genetic engineering, and to the use thereof for drug screening, diagnosis and therapy. A further aspect of the invention also relates to polynucleotides encoding such fusion proteins. Examples of fusion protein technology can be found in International Patent Application Nos. WO94/29458 and WO94/22914.

Each of the polynucleotide sequences provided herein may be used in the discovery and development of antibacterial compounds. The encoded protein, upon expression, can be used as a target for the screening of antibacterial drugs. Additionally, the polynucleotide sequences encoding the amino terminal regions of the encoded protein or Shine-Delgarno or other translation facilitating sequences of the respective mRNA can be used to construct antisense sequences to control the expression of the coding sequence of interest.

The invention also provides the use of the polypeptide, polynucleotide, agonist or antagonist of the invention to interfere with the initial physical interaction between a pathogen or pathogens and a eukaryotic, preferably mammalian, host responsible for sequelae of infection. In particular, the molecules of the invention may be used: in the prevention of adhesion of bacteria, in particular gram positive and/or gram negative bacteria, to eukaryotic, preferably mammalian, extracellular matrix proteins on in-dwelling devices or to extracellular matrix proteins in wounds; to block bacterial adhesion between eukaryotic, preferably mammalian, extracellular matrix proteins and bacterial BASB024 proteins that mediate tissue damage and/or; to block the normal progression of pathogenesis in infections initiated other than by the implantation of in-dwelling devices or by other surgical techniques.

In accordance with yet another aspect of the invention, there are provided BASB024 agonists and antagonists, preferably bacteristatic or bactericidal agonists and antagonists.

The antagonists and agonists of the invention may be employed, for instance, to prevent, inhibit and/or treat diseases.

In a further aspect, the present invention relates to mimotopes of the polypeptide of the invention. A mimotope is a peptide sequence, sufficiently similar to the native peptide (sequentially or structurally), which is capable of being recognised by antibodies which recognise the native peptide; or is capable of raising antibodies which recognise the native peptide when coupled to a suitable carrier.

Peptide mimotopes may be designed for a particular purpose by addition, deletion or substitution of elected amino acids. Thus, the peptides may be modified for the purposes of ease of conjugation to a protein carrier. For example, it may be desirable for some chemical conjugation methods to include a terminal cysteine. In addition it may be desirable for peptides conjugated to a protein carrier to include a hydrophobic terminus

distal from the conjugated terminus of the peptide, such that the free unconjugated end of the peptide remains associated with the surface of the carrier protein. Thereby presenting the peptide in a conformation which most closely resembles that of the peptide as found in the context of the whole native molecule. For example, the peptides may be altered to have an N-terminal cysteine and a C-terminal hydrophobic amidated tail. Alternatively, the addition or substitution of a D-stereoisomer form of one or more of the amino acids may be performed to create a beneficial derivative, for example to enhance stability of the peptide.

Alternatively, peptide mimotopes may be identified using antibodies which are capable themselves of binding to the polypeptides of the present invention using techniques such as phage display technology (EP 0 552 267 B1). This technique, generates a large number of peptide sequences which mimic the structure of the native peptides and are, therefore, capable of binding to anti-native peptide antibodies, but may not necessarily themselves share significant sequence homology to the native polypeptide.

Vaccines

Another aspect of the invention relates to a method for inducing an immunological response in an individual, particularly a mammal, preferably humans, which comprises inoculating the individual with BASB024 polynucleotide and/or polypeptide, or a fragment or variant thereof, adequate to produce antibody and/ or T cell immune response to protect said individual from infection, particularly bacterial infection and most particularly *Neisseria meningitidis* infection. Also provided are methods whereby such immunological response slows bacterial replication. Yet another aspect of the invention relates to a method of inducing immunological response in an individual which comprises delivering to such individual a nucleic acid vector, sequence or ribozyme to direct expression of BASB024 polynucleotide and/or polypeptide, or a fragment or a variant thereof, for expressing BASB024 polynucleotide and/or polypeptide, or a fragment or a variant thereof *in vivo* in order to induce an immunological response, such as, to produce

antibody and/ or T cell immune response, including, for example, cytokine-producing T cells or cytotoxic T cells, to protect said individual, preferably a human, from disease, whether that disease is already established within the individual or not. One example of administering the gene is by accelerating it into the desired cells as a coating on particles or otherwise. Such nucleic acid vector may comprise DNA, RNA, a ribozyme, a modified nucleic acid, a DNA/RNA hybrid, a DNA-protein complex or an RNA-protein complex.

A further aspect of the invention relates to an immunological composition that when introduced into an individual, preferably a human, capable of having induced within it an immunological response, induces an immunological response in such individual to a BASB024 polynucleotide and/or polypeptide encoded therefrom, wherein the composition comprises a recombinant BASB024 polynucleotide and/or polypeptide encoded therefrom and/or comprises DNA and/or RNA which encodes and expresses an antigen of said BASB024 polynucleotide, polypeptide encoded therefrom, or other polypeptide of the invention. The immunological response may be used therapeutically or prophylactically and may take the form of antibody immunity and/or cellular immunity, such as cellular immunity arising from CTL or CD4+ T cells.

A BASB024 polypeptide or a fragment thereof may be fused with co-protein or chemical moiety which may or may not by itself produce antibodies, but which is capable of stabilizing the first protein and producing a fused or modified protein which will have antigenic and/or immunogenic properties, and preferably protective properties. Thus fused recombinant protein, preferably further comprises an antigenic co-protein, such as lipoprotein D from *Haemophilus influenzae*, Glutathione-S-transferase (GST) or beta-galactosidase, or any other relatively large co-protein which solubilizes the protein and facilitates production and purification thereof. Moreover, the co-protein may act as an adjuvant in the sense of providing a generalized stimulation of the immune system of the organism receiving the protein. The co-protein may be attached to either the amino- or carboxy-terminus of the first protein.

Provided by this invention are compositions, particularly vaccine compositions, and methods comprising the polypeptides and/or polynucleotides of the invention and immunostimulatory DNA sequences, such as those described in Sato, Y. *et al.* Science 273: 352 (1996).

Also, provided by this invention are methods using the described polynucleotide or particular fragments thereof, which have been shown to encode non-variable regions of bacterial cell surface proteins, in polynucleotide constructs used in such genetic immunization experiments in animal models of infection with *Neisseria meningitidis*. Such experiments will be particularly useful for identifying protein epitopes able to provoke a prophylactic or therapeutic immune response. It is believed that this approach will allow for the subsequent preparation of monoclonal antibodies of particular value, derived from the requisite organ of the animal successfully resisting or clearing infection, for the development of prophylactic agents or therapeutic treatments of bacterial infection, particularly *Neisseria meningitidis* infection, in mammals, particularly humans.

The invention also includes a vaccine formulation which comprises an immunogenic recombinant polypeptide and/or polynucleotide of the invention together with a suitable carrier, such as a pharmaceutically acceptable carrier. Since the polypeptides and polynucleotides may be broken down in the stomach, each is preferably administered parenterally, including, for example, administration that is subcutaneous, intramuscular, intravenous, or intradermal. Formulations suitable for parenteral administration include aqueous and non-aqueous sterile injection solutions which may contain anti-oxidants, buffers, bacteristatic compounds and solutes which render the formulation isotonic with the bodily fluid, preferably the blood, of the individual; and aqueous and non-aqueous sterile suspensions which may include suspending agents or thickening agents. The formulations may be presented in unit-dose or multi-dose containers, for example, sealed

ampoules and vials and may be stored in a freeze-dried condition requiring only the addition of the sterile liquid carrier immediately prior to use.

The vaccine formulation of the invention may also include adjuvant systems for enhancing the immunogenicity of the formulation. Preferably the adjuvant system raises preferentially a TH1 type of response.

An immune response may be broadly distinguished into two extreme catagories, being a humoral or cell mediated immune responses (traditionally characterised by antibody and cellular effector mechanisms of protection respectively). These categories of response have been termed TH1-type responses (cell-mediated response), and TH2-type immune responses (humoral response).

Extreme TH1-type immune responses may be characterised by the generation of antigen specific, haplotype restricted cytotoxic T lymphocytes, and natural killer cell responses. In mice TH1-type responses are often characterised by the generation of antibodies of the IgG2a subtype, whilst in the human these correspond to IgG1 type antibodies. TH2-type immune responses are characterised by the generation of a broad range of immunoglobulin isotypes including in mice IgG1, IgA, and IgM.

It can be considered that the driving force behind the development of these two types of immune responses are cytokines. High levels of TH1-type cytokines tend to favour the induction of cell mediated immune responses to the given antigen, whilst high levels of TH2-type cytokines tend to favour the induction of humoral immune responses to the antigen.

The distinction of TH1 and TH2-type immune responses is not absolute. In reality an individual will support an immune response which is described as being predominantly TH1 or predominantly TH2. However, it is often convenient to consider the families of

cytokines in terms of that described in murine CD4 +ve T cell clones by Mosmann and Coffman (*Mosmann, T.R. and Coffman, R.L. (1989) TH1 and TH2 cells: different patterns of lymphokine secretion lead to different functional properties. Annual Review of Immunology, 7, p145-173*). Traditionally, TH1-type responses are associated with the production of the INF-γ and IL-2 cytokines by T-lymphocytes. Other cytokines often directly associated with the induction of TH1-type immune responses are not produced by T-cells, such as IL-12. In contrast, TH2- type responses are associated with the secretion of IL-4, IL-5, IL-6 and IL-13.

It is known that certain vaccine adjuvants are particularly suited to the stimulation of either TH1 or TH2 - type cytokine responses. Traditionally the best indicators of the TH1:TH2 balance of the immune response after a vaccination or infection includes direct measurement of the production of TH1 or TH2 cytokines by T lymphocytes *in vitro* after restimulation with antigen, and/or the measurement of the IgG1:IgG2a ratio of antigen specific antibody responses.

Thus, a TH1-type adjuvant is one which preferentially stimulates isolated T-cell populations to produce high levels of TH1-type cytokines when re-stimulated with antigen *in vitro*, and promotes development of both CD8+ cytotoxic T lymphocytes and antigen specific immunoglobulin responses associated with TH1-type isotype.

Adjuvants which are capable of preferential stimulation of the TH1 cell response are described in International Patent Application No. WO 94/00153 and WO 95/17209.

3 De-O-acylated monophosphoryl lipid A (3D-MPL) is one such adjuvant. This is known from GB 2220211 (Ribi). Chemically it is a mixture of 3 De-O-acylated monophosphoryl lipid A with 4, 5 or 6 acylated chains and is manufactured by Ribi Immunochem, Montana. A preferred form of 3 De-O-acylated monophosphoryl lipid A is disclosed in European Patent 0 689 454 B1 (SmithKline Beecham Biologicals SA).

Preferably, the particles of 3D-MPL are small enough to be sterile filtered through a 0.22micron membrane (European Patent number 0 689 454).
3D-MPL will be present in the range of 10μg - 100μg preferably 25-50μg per dose wherein the antigen will typically be present in a range 2-50μg per dose.

Another preferred adjuvant comprises QS21, an Hplc purified non-toxic fraction derived from the bark of Quillaja Saponaria Molina. Optionally this may be admixed with 3 De-O-acylated monophosphoryl lipid A (3D-MPL), optionally together with a carrier.

The method of production of QS21 is disclosed in US patent No. 5,057,540.

Non-reactogenic adjuvant formulations containing QS21 have been described previously (WO 96/33739). Such formulations comprising QS21 and cholesterol have been shown to be successful TH1 stimulating adjuvants when formulated together with an antigen.

Further adjuvants which are preferential stimulators of TH1 cell response include immunomodulatory oligonucleotides, for example unmethylated CpG sequences as disclosed in WO 96/02555.

Combinations of different TH1 stimulating adjuvants, such as those mentioned hereinabove, are also contemplated as providing an adjuvant which is a preferential stimulator of TH1 cell response. For example, QS21 can be formulated together with 3D-MPL. The ratio of QS21 : 3D-MPL will typically be in the order of 1 : 10 to 10 : 1; preferably 1:5 to 5 : 1 and often substantially 1 : 1. The preferred range for optimal synergy is 2.5 : 1 to 1 : 1 3D-MPL: QS21.

Preferably a carrier is also present in the vaccine composition according to the invention. The carrier may be an oil in water emulsion, or an aluminium salt, such as aluminium phosphate or aluminium hydroxide.

A preferred oil-in-water emulsion comprises a metabolisible oil, such as squalene, alpha tocopherol and Tween 80. In a particularly preferred aspect the antigens in the vaccine composition according to the invention are combined with QS21 and 3D-MPL in such an emulsion. Additionally the oil in water emulsion may contain span 85 and/or lecithin and/or tricaprylin.

Typically for human administration QS21 and 3D-MPL will be present in a vaccine in the range of 1μg - 200μg, such as 10-100μg, preferably 10μg - 50μg per dose. Typically the oil in water will comprise from 2 to 10% squalene, from 2 to 10% alpha tocopherol and from 0.3 to 3% tween 80. Preferably the ratio of squalene: alpha tocopherol is equal to or less than 1 as this provides a more stable emulsion. Span 85 may also be present at a level of 1%. In some cases it may be advantageous that the vaccines of the present invention will further contain a stabiliser.

Non-toxic oil in water emulsions preferably contain a non-toxic oil, e.g. squalane or squalene, an emulsifier, e.g. Tween 80, in an aqueous carrier. The aqueous carrier may be, for example, phosphate buffered saline.

A particularly potent adjuvant formulation involving QS21, 3D-MPL and tocopherol in an oil in water emulsion is described in WO 95/17210.

The present invention also provides a polyvalent vaccine composition comprising a vaccine formulation of the invention in combination with other antigens, in particular antigens useful for treating cancers, autoimmune diseases and related conditions. Such a

polyvalent vaccine composition may include a TH-1 inducing adjuvant as hereinbefore described.

While the invention has been described with reference to certain BASB024 polypeptides and polynucleotides, it is to be understood that this covers fragments of the naturally occurring polypeptides and polynucleotides, and similar polypeptides and polynucleotides with additions, deletions or substitutions which do not substantially affect the immunogenic properties of the recombinant polypeptides or polynucleotides.

The antigen can also be delivered in the form of whole bacteria (dead or alive) or as subcellular fractions, these possibilities do include *N.meningitidis* itself.

Compositions, kits and administration

In a further aspect of the invention there are provided compositions comprising a BASB024 polynucleotide and/or a BASB024 polypeptide for administration to a cell or to a multicellular organism.

The invention also relates to compositions comprising a polynucleotide and/or a polypeptide discussed herein or their agonists or antagonists. The polypeptides and polynucleotides of the invention may be employed in combination with a non-sterile or sterile carrier or carriers for use with cells, tissues or organisms, such as a pharmaceutical carrier suitable for administration to an individual. Such compositions comprise, for instance, a media additive or a therapeutically effective amount of a polypeptide and/or polynucleotide of the invention and a pharmaceutically acceptable carrier or excipient. Such carriers may include, but are not limited to, saline, buffered saline, dextrose, water, glycerol, ethanol and combinations thereof. The formulation should suit the mode of administration. The invention further relates to diagnostic and pharmaceutical packs and kits comprising one or more containers filled with one or more of the ingredients of the aforementioned compositions of the invention.

Polypeptides, polynucleotides and other compounds of the invention may be employed alone or in conjunction with other compounds, such as therapeutic compounds.

The pharmaceutical compositions may be administered in any effective, convenient manner including, for instance, administration by topical, oral, anal, vaginal, intravenous, intraperitoneal, intramuscular, subcutaneous, intranasal or intradermal routes among others.

In therapy or as a prophylactic, the active agent may be administered to an individual as an injectable composition, for example as a sterile aqueous dispersion, preferably isotonic.

In a further aspect, the present invention provides for pharmaceutical compositions comprising a therapeutically effective amount of a polypeptide and/or polynucleotide, such as the soluble form of a polypeptide and/or polynucleotide of the present invention, agonist or antagonist peptide or small molecule compound, in combination with a pharmaceutically acceptable carrier or excipient. Such carriers include, but are not limited to, saline, buffered saline, dextrose, water, glycerol, ethanol, and combinations thereof. The invention further relates to pharmaceutical packs and kits comprising one or more containers filled with one or more of the ingredients of the aforementioned compositions of the invention. Polypeptides, polynucleotides and other compounds of the present invention may be employed alone or in conjunction with other compounds, such as therapeutic compounds.

The composition will be adapted to the route of administration, for instance by a systemic or an oral route. Preferred forms of systemic administration include injection, typically by intravenous injection. Other injection routes, such as subcutaneous, intramuscular, or intraperitoneal, can be used. Alternative means for systemic administration include transmucosal and transdermal administration using penetrants such as bile salts or fusidic acids or other detergents. In addition, if a polypeptide or other compounds of the present

invention can be formulated in an enteric or an encapsulated formulation, oral administration may also be possible. Administration of these compounds may also be topical and/or localized, in the form of salves, pastes, gels, solutions, powders and the like.

For administration to mammals, and particularly humans, it is expected that the daily dosage level of the active agent will be from 0.01 mg/kg to 10 mg/kg, typically around 1 mg/kg. The physician in any event will determine the actual dosage which will be most suitable for an individual and will vary with the age, weight and response of the particular individual. The above dosages are exemplary of the average case. There can, of course, be individual instances where higher or lower dosage ranges are merited, and such are within the scope of this invention.

The dosage range required depends on the choice of peptide, the route of administration, the nature of the formulation, the nature of the subject's condition, and the judgment of the attending practitioner. Suitable dosages, however, are in the range of 0.1-100 μg/kg of subject.

A vaccine composition is conveniently in injectable form. Conventional adjuvants may be employed to enhance the immune response. A suitable unit dose for vaccination is 0.5-5 microgram/kg of antigen, and such dose is preferably administered 1-3 times and with an interval of 1-3 weeks. With the indicated dose range, no adverse toxicological effects will be observed with the compounds of the invention which would preclude their administration to suitable individuals.

Wide variations in the needed dosage, however, are to be expected in view of the variety of compounds available and the differing efficiencies of various routes of administration. For example, oral administration would be expected to require higher dosages than administration by intravenous injection. Variations in these dosage levels can be adjusted

using standard empirical routines for optimization, as is well understood in the art.

Sequence Databases, Sequences in a Tangible Medium, and Algorithms

Polynucleotide and polypeptide sequences form a valuable information resource with which to determine their 2- and 3-dimensional structures as well as to identify further sequences of similar homology. These approaches are most easily facilitated by storing the sequence in a computer readable medium and then using the stored data in a known macromolecular structure program or to search a sequence database using well known searching tools, such as the GCG program package.

Also provided by the invention are methods for the analysis of character sequences or strings, particularly genetic sequences or encoded protein sequences. Preferred methods of sequence analysis include, for example, methods of sequence homology analysis, such as identity and similarity analysis, DNA, RNA and protein structure analysis, sequence assembly, cladistic analysis, sequence motif analysis, open reading frame determination, nucleic acid base calling, codon usage analysis, nucleic acid base trimming, and sequencing chromatogram peak analysis.

A computer based method is provided for performing homology identification. This method comprises the steps of: providing a first polynucleotide sequence comprising the sequence of a polynucleotide of the invention in a computer readable medium; and comparing said first polynucleotide sequence to at least one second polynucleotide or polypeptide sequence to identify homology.

A computer based method is also provided for performing homology identification, said method comprising the steps of: providing a first polypeptide sequence comprising the sequence of a polypeptide of the invention in a computer readable medium; and

comparing said first polypeptide sequence to at least one second polynucleotide or polypeptide sequence to identify homology.

All publications and references, including but not limited to patents and patent applications, cited in this specification are herein incorporated by reference in their entirety as if each individual publication or reference were specifically and individually indicated to be incorporated by reference herein as being fully set forth. Any patent application to which this application claims priority is also incorporated by reference herein in its entirety in the manner described above for publications and references.

DEFINITIONS

"Identity," as known in the art, is a relationship between two or more polypeptide sequences or two or more polynucleotide sequences, as the case may be, as determined by comparing the sequences. In the art, "identity" also means the degree of sequence relatedness between polypeptide or polynucleotide sequences, as the case may be, as determined by the match between strings of such sequences. "Identity" can be readily calculated by known methods, including but not limited to those described in (*Computational Molecular Biology*, Lesk, A.M., ed., Oxford University Press, New York, 1988; *Biocomputing: Informatics and Genome Projects*, Smith, D.W., ed., Academic Press, New York, 1993; *Computer Analysis of Sequence Data*, Part I, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; *Sequence Analysis in Molecular Biology*, von Heine, G., Academic Press, 1987; and *Sequence Analysis Primer*, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991; and Carillo, H., and Lipman, D., SIAM *J. Applied Math., 48:* 1073 (1988). Methods to determine identity are designed to give the largest match between the sequences tested. Moreover, methods to determine identity are codified in publicly available computer programs. Computer program methods to determine identity between two sequences include, but are not limited to, the GAP program in the GCG program package (Devereux, J., et al., *Nucleic Acids Research 12(1):*

387 (1984)), BLASTP, BLASTN (Altschul, S.F. et al., *J. Molec. Biol. 215:* 403-410 (1990), and FASTA(Pearson and Lipman Proc. Natl. Acad. Sci. USA 85; 2444-2448 (1988). The BLAST family of programs is publicly available from NCBI and other sources (*BLAST Manual*, Altschul, S., *et al.*, NCBI NLM NIH Bethesda, MD 20894; Altschul, S., *et al.*, *J. Mol. Biol. 215*: 403-410 (1990). The well known Smith Waterman algorithm may also be used to determine identity.

Parameters for polypeptide sequence comparison include the following:
Algorithm: Needleman and Wunsch, J. Mol Biol. 48: 443-453 (1970)
Comparison matrix: BLOSSUM62 from Henikoff and Henikoff,
Proc. Natl. Acad. Sci. USA. 89:10915-10919 (1992)
Gap Penalty: 8
Gap Length Penalty: 2
A program useful with these parameters is publicly available as the "gap" program from Genetics Computer Group, Madison WI. The aforementioned parameters are the default parameters for peptide comparisons (along with no penalty for end gaps).

Parameters for polynucleotide comparison include the following:
Algorithm: Needleman and Wunsch, J. Mol Biol. 48: 443-453 (1970)
Comparison matrix: matches = +10, mismatch = 0
Gap Penalty: 50
Gap Length Penalty: 3
Available as: The "gap" program from Genetics Computer Group, Madison WI. These are the default parameters for nucleic acid comparisons.

A preferred meaning for "identity" for polynucleotides and polypeptides, as the case may be, are provided in (1) and (2) below.

(1) Polynucleotide embodiments further include an isolated polynucleotide comprising a polynucleotide sequence having at least a 50, 60, 70, 80, 85, 90, 95, 97 or 100% identity to the reference sequence of SEQ ID NO:1, wherein said polynucleotide sequence may be identical to the reference sequence of SEQ ID NO:1 or may include up to a certain integer number of nucleotide alterations as compared to the reference sequence, wherein said alterations are selected from the group consisting of at least one nucleotide deletion, substitution, including transition and transversion, or insertion, and wherein said alterations may occur at the 5' or 3' terminal positions of the reference nucleotide sequence or anywhere between those terminal positions, interspersed either individually among the nucleotides in the reference sequence or in one or more contiguous groups within the reference sequence, and wherein said number of nucleotide alterations is determined by multiplying the total number of nucleotides in SEQ ID NO:1 by the integer defining the percent identity divided by 100 and then subtracting that product from said total number of nucleotides in SEQ ID NO:1, or:

$$n_n \leq x_n - (x_n \bullet y),$$

wherein n_n is the number of nucleotide alterations, x_n is the total number of nucleotides in SEQ ID NO:1, y is 0.50 for 50%, 0.60 for 60%, 0.70 for 70%, 0.80 for 80%, 0.85 for 85%, 0.90 for 90%, 0.95 for 95%, 0.97 for 97% or 1.00 for 100%, and $\bullet$ is the symbol for the multiplication operator, and wherein any non-integer product of x_n and y is rounded down to the nearest integer prior to subtracting it from x_n. Alterations of a polynucleotide sequence encoding the polypeptide of SEQ ID NO:2 may create nonsense, missense or frameshift mutations in this coding sequence and thereby alter the polypeptide encoded by the polynucleotide following such alterations.

By way of example, a polynucleotide sequence of the present invention may be identical to the reference sequence of SEQ ID NO:1, that is it may be 100% identical, or it may include up to a certain integer number of nucleic acid alterations as compared to the

reference sequence such that the percent identity is less than 100% identity. Such alterations are selected from the group consisting of at least one nucleic acid deletion, substitution, including transition and transversion, or insertion, and wherein said alterations may occur at the 5' or 3' terminal positions of the reference polynucleotide sequence or anywhere between those terminal positions, interspersed either individually among the nucleic acids in the reference sequence or in one or more contiguous groups within the reference sequence. The number of nucleic acid alterations for a given percent identity is determined by multiplying the total number of nucleic acids in SEQ ID NO:1 by the integer defining the percent identity divided by 100 and then subtracting that product from said total number of nucleic acids in SEQ ID NO:1, or:

$$n_n \leq x_n - (x_n \bullet y),$$

wherein n_n is the number of nucleic acid alterations, x_n is the total number of nucleic acids in SEQ ID NO:1, **y** is, for instance 0.70 for 70%, 0.80 for 80%, 0.85 for 85% etc., • is the symbol for the multiplication operator, and wherein any non-integer product of x_n and **y** is rounded down to the nearest integer prior to subtracting it from x_n.

(2) Polypeptide embodiments further include an isolated polypeptide comprising a polypeptide having at least a 50, 60, 70, 80, 85, 90, 95, 97 or 100% identity to a polypeptide reference sequence of SEQ ID NO:2, wherein said polypeptide sequence may be identical to the reference sequence of SEQ ID NO:2 or may include up to a certain integer number of amino acid alterations as compared to the reference sequence, wherein said alterations are selected from the group consisting of at least one amino acid deletion, substitution, including conservative and non-conservative substitution, or insertion, and wherein said alterations may occur at the amino- or carboxy-terminal positions of the reference polypeptide sequence or anywhere between those terminal positions, interspersed either individually among the amino acids in the reference sequence or in one or more contiguous groups within the reference sequence, and wherein said number of

amino acid alterations is determined by multiplying the total number of amino acids in SEQ ID NO:2 by the integer defining the percent identity divided by 100 and then subtracting that product from said total number of amino acids in SEQ ID NO:2, or:

$$n_a \leq x_a - (x_a \bullet y),$$

wherein n_a is the number of amino acid alterations, x_a is the total number of amino acids in SEQ ID NO:2, y is 0.50 for 50%, 0.60 for 60%, 0.70 for 70%, 0.80 for 80%, 0.85 for 85%, 0.90 for 90%, 0.95 for 95%, 0.97 for 97% or 1.00 for 100%, and • is the symbol for the multiplication operator, and wherein any non-integer product of x_a and y is rounded down to the nearest integer prior to subtracting it from x_a.

By way of example, a polypeptide sequence of the present invention may be identical to the reference sequence of SEQ ID NO:2, that is it may be 100% identical, or it may include up to a certain integer number of amino acid alterations as compared to the reference sequence such that the percent identity is less than 100% identity. Such alterations are selected from the group consisting of at least one amino acid deletion, substitution, including conservative and non-conservative substitution, or insertion, and wherein said alterations may occur at the amino- or carboxy-terminal positions of the reference polypeptide sequence or anywhere between those terminal positions, interspersed either individually among the amino acids in the reference sequence or in one or more contiguous groups within the reference sequence. The number of amino acid alterations for a given % identity is determined by multiplying the total number of amino acids in SEQ ID NO:2 by the integer defining the percent identity divided by 100 and then subtracting that product from said total number of amino acids in SEQ ID NO:2, or:

$$n_a \leq x_a - (x_a \bullet y),$$

wherein n_a is the number of amino acid alterations, x_a is the total number of amino acids in SEQ ID NO:2, y is, for instance 0.70 for 70%, 0.80 for 80%, 0.85 for 85% etc., and • is the symbol for the multiplication operator, and wherein any non-integer product of x_a and y is rounded down to the nearest integer prior to subtracting it from x_a.

"Individual(s)," when used herein with reference to an organism, means a multicellular eukaryote, including, but not limited to a metazoan, a mammal, an ovid, a bovid, a simian, a primate, and a human.

"Isolated" means altered "by the hand of man" from its natural state, *i.e.*, if it occurs in nature, it has been changed or removed from its original environment, or both. For example, a polynucleotide or a polypeptide naturally present in a living organism is not "isolated," but the same polynucleotide or polypeptide separated from the coexisting materials of its natural state is "isolated", as the term is employed herein. Moreover, a polynucleotide or polypeptide that is introduced into an organism by transformation, genetic manipulation or by any other recombinant method is "isolated" even if it is still present in said organism, which organism may be living or non-living.

"Polynucleotide(s)" generally refers to any polyribonucleotide or polydeoxyribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA including single and double-stranded regions.

"Variant" refers to a polynucleotide or polypeptide that differs from a reference polynucleotide or polypeptide, but retains essential properties. A typical variant of a polynucleotide differs in nucleotide sequence from another, reference polynucleotide. Changes in the nucleotide sequence of the variant may or may not alter the amino acid sequence of a polypeptide encoded by the reference polynucleotide. Nucleotide changes may result in amino acid substitutions, additions, deletions, fusions and truncations in the polypeptide encoded by the reference sequence, as discussed below. A typical

variant of a polypeptide differs in amino acid sequence from another, reference polypeptide. Generally, differences are limited so that the sequences of the reference polypeptide and the variant are closely similar overall and, in many regions, identical. A variant and reference polypeptide may differ in amino acid sequence by one or more substitutions, additions, deletions in any combination. A substituted or inserted amino acid residue may or may not be one encoded by the genetic code. A variant of a polynucleotide or polypeptide may be a naturally occurring such as an allelic variant, or it may be a variant that is not known to occur naturally. Non-naturally occurring variants of polynucleotides and polypeptides may be made by mutagenesis techniques or by direct synthesis.

"Disease(s)" means any disease caused by or related to infection by a bacteria, including , for example, upper respiratory tract infection, invasive bacterial diseases, such as bacteremia and meningitis.

EXAMPLES:

The examples below are carried out using standard techniques, which are well known and routine to those of skill in the art, except where otherwise described in detail. The examples are illustrative, but do not limit the invention.

Example 1: Discovery and confirmatory DNA sequencing of the BASB024 gene from two *N.meningitidis* strains.

A: BASB024 in *N. meningitidis* serogroup B strain ATCC13090.

The BASB024 gene of SEQ ID NO:1 was first discovered in the Incyte PathoSeq database containing unfinished genomic DNA sequences of the *N. meningitidis* strain ATCC13090. The translation of the BASB024 polynucleotide sequence, showed in SEQ ID NO:2, showed significant similarity (24 % identity in a 837 amino acids overlap) to the *Serratia marcescens* HasR heme acquisition outer membrane protein. The sequence of the BASB024 gene was further confirmed experimentally. For this purpose, genomic DNA was extracted from 10^{10} cells of the *N.meningitidis* cells (strain ATCC 13090) using the QIAGEN genomic DNA extraction kit (Qiagen Gmbh), and 1μg of this material was submitted to Polymerase Chain Reaction DNA amplification using primers HasR01 (5'- GCT AGC TAG CAG ATC TTC TTT CGG GAA GAA -3') [SEQ ID NO:7] containing an internal *Nhe*I site and HasR02 (5'- GGT CGC TCG AGA AAC TTG TAG CTC ATC GTT ATC A -3') [SEQ ID NO:8] containing an internal *Xho*I site. This PCR product was gel-purified and subjected to DNA sequencing using the Big Dye Cycle Sequencing kit (Perkin-Elmer) and an ABI 373A/PRISM DNA sequencer. DNA sequencing was performed on both strands with a redundancy of 2 and the full-length sequence was assembled using the SeqMan program from the DNASTAR Lasergene software package. The resulting DNA sequence and deduced polypeptide sequence are shown as SEQ ID NO:3 and SEQ ID NO:4 respectively.

B: BASB024 in *N. meningitidis* serogroup B strain H44/76.

The sequence of the BASB024 gene was also determined in another *N. meningitidis* serogroup B strain, the strain H44/76. For this purpose, genomic DNA was extracted from the *N. meningitidis* strain H44/76 using the experimental conditions presented in Example 1. This material (1µg) was then submitted to Polymerase Chain Reaction DNA amplification using primers HasR01 and HasR02 specific for the BASB024 gene. The resulting DNA fragment was obtained, digested by the *NheI/XhoI* restriction endonucleases and inserted into the corresponding sites of the pET-24b cloning/expression vector (Novagen) using standard molecular biology techniques (Molecular Cloning,a Laboratory Manual, Second Edition, Eds: Sambrook, Fritsch & Maniatis, Cold Spring Harbor press 1989). Recombinant pET-24b/BASB024 was then submitted to DNA sequencing using the Big Dyes kit (Applied biosystems) and analyzed on a ABI 373/A DNA sequencer in the conditions described by the supplier.

As a result, the polynucleotide and deduced polypeptide sequences, referred to as SEQ ID NO:5 and SEQ ID NO:6 respectively, were obtained. Using the MegAlign program from the DNASTAR software package, an alignment of the polynucleotide sequences of SEQ ID NO:1, 3 and 5 was performed, and is displayed in Figure 1; a pairwise comparison of identities is summarized in Table 1, showing that the three BASB024 polynucleotide gene sequences are all similar at identity level greater than 98.0 %. Using the same MegAlign program, an alignment of the polypeptide sequences of SEQ ID NO:2, 4 and 6 was performed, and is displayed in Figure 2; a pairwise comparison of identities is summarized in Table 2, showing that the three BASB024 protein sequences are all similar at a identity level greater than 97.0 %.

Taken together, these data indicate strong sequence conservation of the BASB024 gene among the two *N.meningitidis* serogroup B strains.

Table 1: Pairwise identities of the BASB024 polynucleotide sequences (in %)

	SeqID No:3	SeqID No:5
SeqID No:1	99.8	98.0
SeqID No:3		98.2

Table 2: Pairwise identities of the BASB024 polypeptide sequences (in %)

	SeqID No:4	SeqID No:6
SeqID No:2	99.8	97.7
SeqID No:4		97.9

Example 2: Expression and purification of recombinant BASB024 protein in *Escherichia coli.*

The construction of the pET-24b/BASB024 cloning/expression vector was described in Example 1B. This vector harbours the BASB024 gene isolated from the strain H44/76 in fusion with a stretch of 6 Histidine residues, placed under the control of the strong bacteriophage T7 gene 10 promoter. For expression study, this vector was introduced into the *Escherichia coli* strain Novablue (DE3) (Novagen), in which, the gene for the T7 polymerase is placed under the control of the isopropyl-beta-D thiogalactoside (IPTG)-regulatable *lac* promoter. Liquid cultures (100 ml) of the Novablue (DE3) [pET-24b/BASB024] *E. coli* recombinant strain were grown at 37°C under agitation until the optical density at 600nm (OD600) reached 0.6. At that time-point, IPTG was added at a final concentration of 1mM and the culture was grown for 4 additional hours. The

culture was then centrifuged at 10,000 rpm and the pellet was frozen at -20°C for at least 10 hours. As presented in Figure 3, SDS PAGE and immuno-blotting (anti-His mAb) analysis of induced bacterial protein extracts deteced an IPTG-inducible recombinant protein of ~ 85kDa.

Legend for Figure 3: Substantially pure (more than 80%) BASB024 protein fractions were obtained on a 4-20% gradient polyacrylamide gel (NOVEX) under SDS-PAGE conditions in parallel to a protein molecular weight marker. Gels were either stained with Coomassie Blue R250 or analyzed by western blot using an anti-(His5) monoclonal antibody. (Non-in. stands for non–induced bacterial cultures, In. stands for induced bacterial cultures).

After thawing, the pellet (2 litre culture) was resuspended during 30 minutes at 22°C in 20 mM phosphate buffer pH 7.0 containing 1 mM Pefabloc prior cell lysis by three passes through a French press disruptor. Lysed cells were pelleted 30 min at 15,000 rpm (Beckman J2-HS centrifuge, JA-20 rotor) at 4°C.

BASB024/His6 was solubilised by 6 M Guanidine Chloride,1 mM Pefabloc, 20 mM phosphate pH 8.0 during 4 hours at 4°C. Cell debris were pelleted 30 min at 15,000 rpm in a JA-20 rotor at 4°C. 0.5 M NaCl was added to the sample before chromatography.

The sample was loaded at a flow-rate of 0.5 ml/min on a 2 ml Ni2+-NTA Agarose column (Qiagen). The column was equilibrated in 6 M Guanidine Chloride, 500 mM NaCl, 20 mM phosphate pH 8.0.

After passage of the flowthrough, the column was washed with equilibration buffer until the base line was reached and then with 10 column volumes of 20 mM phosphate pH 8.0, 6M Guanidine Chloride, 5 mM Imidazole. The recombinant protein was eluted from the column by 100 mM Imidazole in 6M Guanidine Chloride, 20 mM phosphate pH 8.0, at 1 ml/min.

IMAC-eluted sample was dialysed at 4°C versus PBS containing 0.5 M Arginine.

As shown in Figure 4 (lane 5), an enriched (purity estimated at 5% pure in CBB stained SDS-PAGE) BASB024 protein, migrating at 105kDa (estimated relative molecular

mass), was eluted from the column. This polypeptide was reactive against a mouse monoclonal antibody raised against the 5-histidine motif (see Figure 5).
Taken together, these data indicate that the BASB024 gene can be expressed and purified under a recombinant form (BASB024/His6) in *E.coli*.

Example 3: Presence of anti-BASB024 antibodies in sera from convalescent patients.

In this test, several convalescent sera have been tested by western-blotting for recognition of the purified recombinant BASB024 protein.
Briefly, around 10 µg of partially purified BASB024 *N.meningitidis* serogroup B protein are put into a SDS-PAGE gradient gel (4-20%, Novex, code n°EC6028) for electrophoretic migration. Proteins are transferred to nitrocellulose sheet (0.45 µm, Bio-rad code n° 162-0114) at 100 volts for 1 hour using a Bio-rad Trans-blot system (code n°170-3930). Afterwards, filter is blocked with PBS - 0.05 % Tween 20 overnight at room temperature, before incubation with the human sera (these sera are diluted 100 times in PBS - 0.05 % Tween 20) or with anti-pentaHis mouse antibody (Qiagen 34660) (this antibody is diluted 200 times in PBS - 0.05 % Tween 20, and incubated on the nitrocellulose sheet for two hours at room temperature with gentle shaking. After three repeated washing steps in PBS - 0.05 % Tween 20 for 5 min., the nitrocellulose sheet is incubated at room temperature for 1 hour under gentle shaking with the appropriate conjugate (biotinylated anti-human Ig antibodies, from sheep, Amersham code n°RPN1003 or anti-mouse Ig antibodies, Amersham code n°RPN1001) diluted at 1/500 in the same washing buffer. The membrane is washed three times as previously, and incubated for 30 min with agitation using the streptavidin-peroxidase complex (Amersham code n°1051) diluted at 1/1000 in the washing buffer. After the last three repeated washing steps, the revelation occurs during the 20 min incubation time in a 50 ml solution containing 30 mg 4-chloro-1-naphtol (Sigma), 10 ml methanol,

40 ml of ultra-pure water, and 30 µl of H_2O_2. The staining is stopped while washing the membrane several times in distillated water.

Results illustrated hereafter in Figure 6 show that all the 3 convalescents react against a lot of *E.coli* contaminants, the BASB024 protein being around 10 % pure. However, the recombinant BASB024 protein is clearly visible with the 3 convalescents at around 105 kDa, meaning that all of them recognize this protein after a *Neisseria meningitidis* B infection.. The anti-His tail monoclonal antibody used clearly localize the BASB024 protein on the gel.

Example 4 : Analysis of the non-coding flanking regions of the BASB024 gene, and its exploitation for modulated BASB024 gene expression.

The non-coding flanking regions of the BASB024 gene contain regulatory elements important in the expression of the gene. This regulation takes place both at the transcriptional and translational level. The sequence of these regions, either upstream or downstream of the open reading frame of the gene, can be obtained by DNA sequencing. This sequence information allows the determination of potential regulatory motifs such as the different promoter elements, terminator sequences, inducible sequence elements, repressors, elements responsible for phase variation, the shine-dalgarno sequence, regions with potential secondary structure involved in regulation, as well as other types of regulatory motifs or sequences.

This sequence information allows the modulation of the natural expression of gene BASB024. The upregulation of the gene expression may be accomplished by altering the promoter, the shine-dalgarno sequence, potential repressor or operator elements, or any other elements involved. Likewise, downregulation of expression can be achieved by similar types of modifications. Alternatively, by changing phase variation sequences, the expression of the gene can be put under phase variation control, or may be uncoupled from this regulation. In another approach, the expression of the gene can be put under the control of

one or more inducible elements allowing regulated expression. Examples of such regulation include, but are not limited to, induction by temperature shift, addition of inductor substrates like selected carbohydrates or their derivatives, trace elements, vitamins, co-factors, metal ions, etc.

Such modifications as described above can be introduced by several different means. The modification of sequences involved in gene expression can be done *in vivo* by random mutagenesis followed by selection for the desired phenotype. Another approach consists in isolating the region of interest and modifying it by random mutagenesis, or site-directed mutagenesis, insertion or deletion mutagenesis. The modified region can then be reintroduced into the bacterial genome by homologous recombination, and the effect on gene expression can be assessed. In another approach, the sequence knowledge of the region of interest can be used to replace or delete all or part of the natural regulatory sequences. In this case, the regulatory region targeted is isolated and modified so as to contain the regulatory elements from another gene, a combination of regulatory elements from different genes, a synthetic regulatory region, or any other regulatory region, or to delete selected parts of the wild-type regulatory sequences. These modified sequences can then be reintroduced into the bacterium via homologous recombination into the genome. A non-exhaustive list of preferred promoters that could be used for up-regulation of gene expression includes the promoter *porA, porB, lbpB, tbpB, p110, lst, hpuAB* from *N. meningitidis* or *N. gonorroheae.*

In one example, the expression of the gene can be modulated by exchanging its promoter with a stronger promoter (through isolating the upstream sequence of the gene, in vitro modification of this sequence, and reintroduction into the genome by homologous recombination). Upregulated expression can be obtained in both the bacterium as well as in the outer membrane vesicles shed (or made) from the bacterium.

In other examples, the described approaches can be used to generate recombinant bacterial strains with improved characteristics for vaccine applications. These can be, but are not limited to, attenuated strains, strains with increased expression of selected antigens, strains with knock-outs (or decreased expression) of genes interfering with the immune response,

strains with modulated expression of immunodominant proteins, strains with modulated shedding of outer-membrane vesicles.

A region directly upstream of the BASB024 gene is given in the sequence of SEQ ID NO:9. This sequence is a further aspect of the invention.

Deposited materials

A deposit containing a *Neisseria meningitidis* Serogroup B strain has been deposited with the American Type Culture Collection (herein "ATCC") on June 22, 1997 and assigned deposit number 13090. The deposit was described as *Neisseria meningitidis* (Albrecht and Ghon) and is a freeze-dried, 1.5-2.9 kb insert library constructed from *N. meningitidis* isolate. The deposit is described in Int. Bull. Bacteriol. Nomencl. Taxon. 8: 1-15 (1958).

The *Neisseria meningitidis* strain deposit is referred to herein as "the deposited strain" or as "the DNA of the deposited strain."

The deposited strain contains the full length BASB024 gene. The sequence of the polynucleotides contained in the deposited strain, as well as the amino acid sequence of any polypeptide encoded thereby, are controlling in the event of any conflict with any description of sequences herein.

The deposit of the deposited strain has been made under the terms of the Budapest Treaty on the International Recognition of the Deposit of Micro-organisms for Purposes of Patent Procedure. The strain will be irrevocably and without restriction or condition released to the public upon the issuance of a patent. The deposited strain is provided merely as convenience to those of skill in the art and is not an admission that a deposit is required for enablement, such as that required under 35 U.S.C. §112.

Applicant's or agent's file reference	KP/BM45330	International application No.

INDICATIONS RELATING TO DEPOSITED MICROORGANISM OR OTHER BIOLOGICAL MATERIAL

(PCT Rule 13*bis*)

A. The indications made below relate to the deposited microorganism or other biological material referred to in the description on page 57, line 3-22

B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet ☐

Name of depositary institution

AMERICAN TYPE CULTURE COLLECTION

Address of depositary institution *(including postal code and country)*

10801 UNIVERSITY BLVD, MANASSAS, VIRGINIA 20110-2209,
UNITED STATES OF AMERICA

Date of deposit	Accession Number
22 June 1997 (22.06.1997)	13090

C. ADDITIONAL INDICATIONS *(leave blank if not applicable)* This information is continued on an additional sheet ☐

In respect of those designations where a European Patent is sought, a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European Patent or until the date on which the application has been refused or withdrawn, only by issue of such a sample to an expert nominated by the person requesting the sample.

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE *(if the indications are not for all designated States)*

E. SEPARATE FURNISHING OF INDICATIONS *(leave blank if not applicable)*

The indications listed below will be submitted to the International Bureau later *(specify the general nature of the indications e.g., "Accession Number of Deposit")*

For receiving Office use only	For International Bureau use only
☐ This sheet was received with the international application	☐ This sheet was received by the International Bureau on:
(13 08 99) 13 aug. 1999	
Authorized officer Y. Marinus-v.d. Nouweland	Authorized officer

CLAIMS:

1. An isolated polypeptide comprising an amino acid sequence which has at least 85% identity to the amino acid sequence selected from the group consisting of: SEQ ID NO:4 and SEQ ID NO:6.

2. An isolated polypeptide as claimed in claim 1 in which the amino acid sequence has at least 95% identity to the amino acid sequence selected from the group consisting of: SEQ ID NO:4 and SEQ ID NO:6.

3. The polypeptide as claimed in claim 1 comprising the amino acid sequence selected from the group consisting of: SEQ ID NO:4 and SEQ ID NO:6.

4. An isolated polypeptide of SEQ ID NO:4 and SEQ ID NO:6.

5. An isolated polypeptide of SEQ ID NO:2.

6. An immunogenic fragment of the polypeptide as claimed in any one of claims 1 to 5 in which the immunogenic activity of said immunogenic fragment is substantially the same as the polypeptide of SEQ ID NO:4 or SEQ ID NO:6.

7. An isolated polynucleotide comprising a nucleotide sequence encoding a polypeptide that has at least 85% identity to the amino acid sequence of SEQ ID NO: 4 or 6 over the entire length of SEQ ID NO: 4 or 6 respectively; or a nucleotide sequence complementary to said isolated polynucleotide.

8. An isolated polynucleotide comprising a nucleotide sequence that has at least 85% identity to a nucleotide sequence encoding a polypeptide of SEQ ID NO: 4 or 6 over the

entire coding region; or a nucleotide sequence complementary to said isolated polynucleotide.

9. An isolated polynucleotide which comprises a nucleotide sequence which has at least 85% identity to that of SEQ ID NO: 3 or 5 over the entire length of SEQ ID NO: 3 or 5 respectively; or a nucleotide sequence complementary to said isolated polynucleotide.

10. The isolated polynucleotide as claimed in any one of claims 7 to 9 in which the identity is at least 95% to SEQ ID NO: 3 or 5.

11. An isolated polynucleotide comprising a nucleotide sequence encoding the polypeptide of SEQ ID NO:4 or SEQ ID NO:6.

12. An isolated polynucleotide comprising the polynucleotide of SEQ ID NO:3 or SEQ ID NO:5.

13. An isolated polynucleotide comprising a nucleotide sequence encoding the polypeptide of SEQ ID NO:4 or SEQ ID NO:6, obtainable by screening an appropriate library under stringent hybridization conditions with a labeled probe having the sequence of SEQ ID NO:3 or SEQ ID NO:5 or a fragment thereof.

14. An isolated polynucleotide comprising a nucleotide sequence encoding the polypeptide of SEQ ID NO:2.

15. An isolated polynucleotide comprising the polynucleotide of SEQ ID NO:1.

16. An isolated polynucleotide comprising a nucleotide sequence encoding the polypeptide of SEQ ID NO:2, obtainable by screening an appropriate library under stringent

hybridization conditions with a labeled probe having the sequence of SEQ ID NO:1 or a fragment thereof.

17. An expression vector or a recombinant live microorganism comprising an isolated polynucleotide according to any one of claims 7 - 16.

18. A host cell comprising the expression vector of claim 17 or a subcellular fraction or a membrane of said host cell expressing an isolated polypeptide comprising an amino acid sequence that has at least 85% identity to the amino acid sequence selected from the group consisting of: SEQ ID NO:4 or SEQ ID NO:6.

19. A process for producing a polypeptide comprising an amino acid sequence that has at least 85% identity to the amino acid sequence selected from the group consisting of SEQ ID NO:4 or SEQ ID NO:6 comprising culturing a host cell of claim 18 under conditions sufficient for the production of said polypeptide and recovering the polypeptide from the culture medium.

20. A process for expressing a polynucleotide of any one of claims 7 – 16 comprising transforming a host cell with the expression vector comprising at least one of said polynucleotides and culturing said host cell under conditions sufficient for expression of any one of said polynucleotides.

21. A vaccine composition comprising an effective amount of the polypeptide of any one of claims 1 to 6 and a pharmaceutically acceptable carrier.

22. A vaccine composition comprising an effective amount of the polynucleotide of any one of claims 7 to 16 and a pharmaceutically effective carrier.

23. The vaccine composition according to either one of claims 21 or 22 wherein said composition comprises at least one other *Neisseria meningitidis* antigen.

24. An antibody immunospecific for the polypeptide or immunological fragment as claimed in any one of claims 1 to 6.

25. A method of diagnosing a *Neisseria meningitidis* infection, comprising identifying a polypeptide as claimed in any one of claims 1 - 6, or an antibody that is immunospecific for said polypeptide, present within a biological sample from an animal suspected of having such an infection.

26. Use of a composition comprising an immunologically effective amount of a polypeptide as claimed in any one of claims 1 – 6 in the preparation of a medicament for use in generating an immune response in an animal.

27. Use of a composition comprising an immunologically effective amount of a polynucleotide as claimed in any one of claims 7 - 16 in the preparation of a medicament for use in generating an immune response in an animal.

28. A therapeutic composition useful in treating humans with *Neisseria meningitidis* disease comprising at least one antibody directed against the polypeptide of claims 1 – 6 and a suitable pharmaceutical carrier.

Figure 1 : Alignment of the BASB024 polynucleotide sequences.
Identity to SeqID No:1 is indicated by a dot, and a dash ("-") indicates a missing nucleotide.

```
                 *         20          *         40          *
Seqid1 : ATGAGATCTTCTTTCCGGTTGAAGCCGATTTGTTTTTATCTTATGGGTGT :   50
Seqid3 : ..................................................... :   50
Seqid5 : .....................................C.............. :   50

                60          *         80          *        100
Seqid1 : TATGCTATATCATCATAGTTATGCCGAAGATGCAGGGCGCGCGGGCAGCG :  100
Seqid3 : ..................................................... :  100
Seqid5 : ..C.........T........................................ :  100

                 *        120          *        140          *
Seqid1 : AGGCGCAGATACAGGTTTTGGAAGATGTGCACGTCAAGGCGAAGCGCGTA :  150
Seqid3 : ..................................................... :  150
Seqid5 : ..................................................... :  150

               160          *        180          *        200
Seqid1 : CCGAAAGACAAAAAAGTGTTTACCGATGCGCGTGCCGTATCGACCCGTCA :  200
Seqid3 : ..................................................... :  200
Seqid5 : ..................................................... :  200

                 *        220          *        240          *
Seqid1 : GGATATATTCAAATCCAGCGAAAACCTCGACAACATCGTACGCAGCATCC :  250
Seqid3 : ..................................................... :  250
Seqid5 : ..................................................... :  250

               260          *        280          *        300
Seqid1 : CCGGTGCGTTTACACAGCAAGATAAAAGCTCGGGCATTGTGTCTTTGAAT :  300
Seqid3 : ..................................................... :  300
Seqid5 : ..................................................... :  300
```

```
                  *       320         *       340         *
Seqid1 : ATTCGCGGCGACAGCGGGTTCGGGCGGGTCAATACGATGGTGGACGGCAT :  350
Seqid3 : ..................................................... :  350
Seqid5 : ..................................................... :  350

                  360         *       380         *       400
Seqid1 : CACGCAGACCTTTTATTCGACTTCTACCGATGCGGGCAGGGCAGGCGGTT :  400
Seqid3 : ..................................................... :  400
Seqid5 : ..................................................... :  400

                  *       420         *       440         *
Seqid1 : CATCTCAATTCGGTGCATCTGTCGACAGCAATTTTATTGCCGGACTGGAT :  450
Seqid3 : ..................................................... :  450
Seqid5 : ..................................................... :  450

                  460         *       480         *       500
Seqid1 : GTCGTCAAAGGCAGCTTCAGCGGCTCGGCAGGCATCAACAGCCTTGCCGG :  500
Seqid3 : ..................................................... :  500
Seqid5 : ..................................................... :  500

                  *       520         *       540         *
Seqid1 : TTCGGCGAATCTGCGGACTTTAGGCGTGGATGACGTCGTTCAGGGCAATA :  550
Seqid3 : ..................................................... :  550
Seqid5 : ..................................................... :  550

                  560         *       580         *       600
Seqid1 : ATACCTACGGCCTGCTGCTAAAAGGTCTGACCGGCACCAATTCAACCAAA :  600
Seqid3 : ..................................................... :  600
Seqid5 : ..................................................... :  600

                  *       620         *       640         *
Seqid1 : GGTAATGCGATGGCGGCGATAGGTGCGCGCAAATGGCTGGAAAGCGGAGC :  650
```

```
Seqid3 : ..................................................... :  650
Seqid5 : ..................................................... :  650


                660         *       680         *       700
Seqid1 : ATCTGTCGGTGTGCTTTACGGGCACAGCAGGCGCACGTGGGCGCAAAATT :  700
Seqid3 : ..................................GCGT.............. :  700
Seqid5 : ..................................GCGT.............. :  700


                    *       720         *       740         *
Seqid1 : ACCGCGTGGGCGGCGGGCAGCACATCGGAAATTTTGGCGCGGAATAT :  750
Seqid3 : ..................................................... :  750
Seqid5 : ..................................................... :  750


                760         *       780         *       800
Seqid1 : CTGGAACGGCGCAAACAGCGATATTTTGTACAAGAAGGCGGGTTGAAATT :  800
Seqid3 : ..............G...................................... :  800
Seqid5 : T.............G.....................G..T.CT........ :  800


                    *       820         *       840         *
Seqid1 : CAATTCCAACAGCGGAAAATGGGAGCGGGATTTCCAAAGGCCGTACTGGA :  850
Seqid3 : ..................................................... :  850
Seqid5 : .......G........................A.......AAC.G.... :  850


                860         *       880         *       900
Seqid1 : AAACCAAGTGGTATCAAAAATACAATGACCCCCAAGAACTGCAAAAATAC :  900
Seqid3 : ..................................................... :  900
Seqid5 : ..TA....CC....A....T.....CA...---.......A......... :  897


                    *       920         *       940         *
Seqid1 : ATCGAAGGTCATGACAAAAGCTGGCGGGAAAACCTGGCGCCGCAATACGA :  950
Seqid3 : ..................................................... :  950
Seqid5 : ..................................................... :  947
```

```
                    960         *        980         *       1000
Seqid1 : CATCACCCCCATCGATCCGTCCAGCCTGAAGCAGCAGTCGGCAGGCAATC : 1000
Seqid3 : .................................................. : 1000
Seqid5 : .................................................. :  997

                *        1020        *        1040        *
Seqid1 : TGTTTAAATTGGAATACGACGGCGTATTCAATAAATACACGGCGCAATTT : 1050
Seqid3 : .................................................. : 1050
Seqid5 : .................................................. : 1047

                   1060         *        1080        *       1100
Seqid1 : CGCGATTTAAACACCAAAATCGGCAGCCGCAAAATCATCAACCGCAATTA : 1100
Seqid3 : .................................................. : 1100
Seqid5 : .................................................. : 1097

                *        1120        *        1140        *
Seqid1 : TCAATTCAATTACGGTTTATCTTTAAACTCATATGCCAACCTCAATCTGA : 1150
Seqid3 : .................................................. : 1150
Seqid5 : ...G..............G.....G...C.G...A............... : 1147

                   1160         *        1180        *       1200
Seqid1 : CCGCAGCCTACAATTCGGGCAGGCAGAAATATCCGAAAGGGTCGAAGTTT : 1200
Seqid3 : .................................................. : 1200
Seqid5 : .................................................. : 1197

                *        1220        *        1240        *
Seqid1 : ACAGGCTGGGGGCTTTTAAAAGATTTTGAAACCTACAACAACGCGAAAAT : 1250
Seqid3 : .................................................. : 1250
Seqid5 : ....................G............................. : 1247

                   1260         *        1280        *       1300
Seqid1 : CCTCGACCTCAACAACACCGCCACCTTCCGGCTGCCCCGCGAAACCGAGT : 1300
Seqid3 : .................................................. : 1300
Seqid5 : .................................................. : 1297
```

```
                    *      1320          *      1340          *
Seqid1 : TGCAAACCACTTTGGGCTTCAATTATTTCCACAACGAATACGGCAAAAAC : 1350
Seqid3 : ..................................................... : 1350
Seqid5 : ..................................................... : 1347


                1360          *      1380          *      1400
Seqid1 : CGCTTTCCTGAAGAATTGGGGCTGTTTTTCGACGGTCCGGATCAGGACAA : 1400
Seqid3 : ..................................................... : 1400
Seqid5 : .....................................T............ : 1397


                    *      1420          *      1440          *
Seqid1 : CGGGCTTTATTCCTATTTGGGGCGGTTTAAGGGCGATAAAGGGCTGCTGC : 1450
Seqid3 : ..................................................... : 1450
Seqid5 : ..................................................... : 1447


                1460          *      1480          *      1500
Seqid1 : CCCAAAAATCAACCATCGTCCAACCGGCCGGCAGCCAATATTTCAACACG : 1500
Seqid3 : ..................................................... : 1500
Seqid5 : ................T.............................. : 1497


                    *      1520          *      1540          *
Seqid1 : TTCTACTTCGATGCCGCGCTCAAAAAAGACATTTACCGCTTAAACTACAG : 1550
Seqid3 : ..................................................... : 1550
Seqid5 : ..................................................... : 1547


                1560          *      1580          *      1600
Seqid1 : CACCAATACCGTCGGCTACCGTTTCGGCGGCGAATATACGGGCTATTACG : 1600
Seqid3 : ..................................................... : 1600
Seqid5 : ..................................................... : 1597


                    *      1620          *      1640          *
Seqid1 : GCTCGGATGACGAATTTAAGCGGGCATTCGGAGAAAACTCGCCGACATAC : 1650
```

```
Seqid3 : ..................................................... : 1650
Seqid5 : ..................................................... : 1647


                1660          *       1680          *       1700
Seqid1 : AAGAAACATTGCAACCAGAGCTGCGGAATTTATGAACCCGTATTGAAAAA : 1700
Seqid3 : ..................................................... : 1700
Seqid5 : ................G.........G..................... : 1697


                    *       1720          *       1740          *
Seqid1 : ATACGGCAAAAAGCGCGCCAACAACCATTCGGTCAGCATTAGTGCGGACT : 1750
Seqid3 : ..................................................... : 1750
Seqid5 : ..................................................... : 1747


                1760          *       1780          *       1800
Seqid1 : TCGGCGATTATTTCATGCCGTTCGCCAGCTATTCGCGCACACACCGTATG : 1800
Seqid3 : ..................................................... : 1800
Seqid5 : ..................................................... : 1797


                    *       1820          *       1840          *
Seqid1 : CCCAACATCCAAGAAATGTATTTTTCCCAAATCGGCGACTCCGGCGTTCA : 1850
Seqid3 : ..................................................... : 1850
Seqid5 : ..................................................... : 1847


                1860          *       1880          *       1900
Seqid1 : CACCGCCTTAAAACCAGAGCGCGCAAACACTTGGCAATTTGGCTTCAATA : 1900
Seqid3 : ..................................................... : 1900
Seqid5 : ..................................................... : 1897


                    *       1920          *       1940          *
Seqid1 : CCTATAAAAAAGGATTGTTAAAACAAGATGATACATTAGGATTAAAACTG : 1950
Seqid3 : ..................................................... : 1950
Seqid5 : ..................................................... : 1947
```

```
                1960          *       1980          *       2000
Seqid1 : GTCGGCTACCGCAGCCGCATCGACAACTACATCCACAACGTTTACGGGAA : 2000
Seqid3 : ..................................................... : 2000
Seqid5 : ..................................................... : 1997

                      *       2020          *       2040          *
Seqid1 : ATGGTGGGATTTGAACGGGAATATTCCGAGCTGGGTCAGCAGCACCGGGC : 2050
Seqid3 : ..................................................... : 2050
Seqid5 : ..................G.............................. : 2047

                2060          *       2080          *       2100
Seqid1 : TTGCCTACACCATCCAACACCGCAATTTCAAAGACAAAGTACACAAACAC : 2100
Seqid3 : ..................................................... : 2100
Seqid5 : ..................T.....................G......... : 2097

                      *       2120          *       2140          *
Seqid1 : GGTTTTGAGTTGGAGCTGAATTACGATTATGGGCGTTTTTTCACCAACCT : 2150
Seqid3 : ..................................................... : 2150
Seqid5 : ..................................................... : 2147

                2160          *       2180          *       2200
Seqid1 : TTCTTACGCCTATCAAAAAAGCACGCAACCGACCAACTTCAGCGATGCGA : 2200
Seqid3 : ..................................................... : 2200
Seqid5 : ..................................................... : 2197

                      *       2220          *       2240          *
Seqid1 : GCGAATCGCCCAACAATGCGTCCAAAGAAGACCAACTCAAACAAGGTTAT : 2250
Seqid3 : ..................................................... : 2250
Seqid5 : ..................................................... : 2247

                2260          *       2280          *       2300
Seqid1 : GGGTTGAGCAGGGTTTCCGCCCTGCCGCGAGATTACGGACGTTTGGAAGT : 2300
Seqid3 : ..................................................... : 2300
```

```
Seqid5 : ..................................................... : 2297


                    *      2320          *      2340          *
Seqid1 : CGGTACGCGCTGGTTGGGCAACAAACTGACTTTGGGCGGCGCGATGCGCT : 2350
Seqid3 : ..................................................... : 2350
Seqid5 : ..................................................... : 2347


               2360          *      2380          *      2400
Seqid1 : ATTTCGGCAAGAGCATCCGCGCGACGGCTGAAGAACGCTATATCGACGGC : 2400
Seqid3 : ..................................................... : 2400
Seqid5 : ..................................................... : 2397


                    *      2420          *      2440          *
Seqid1 : ACCAACGGGGGAAATACCAGCAATGTCCGGCAACTGGGCAAGCGTTCCAT : 2450
Seqid3 : ..................................................... : 2450
Seqid5 : .....................T................................ : 2447


               2460          *      2480          *      2500
Seqid1 : CAAACAAACCGAAACCCTTGCCCGCCAGCCTTTGATTTTTGATTTTTACG : 2500
Seqid3 : ..................................................... : 2500
Seqid5 : ..............T...................................... : 2497


                    *      2520          *      2540          *
Seqid1 : CCGCTTACGAGCCGAAGAAAAACCTTATTTCCGCGCCGAAGTCAAAAAT : 2550
Seqid3 : ..................................................... : 2550
Seqid5 : ..................................................... : 2547


               2560          *      2580          *      2600
Seqid1 : CTGTTCGACAGGCGTTATATCGATCCGCTCGATGCGGGCAATGATGCGGC : 2600
Seqid3 : ..................................................... : 2600
Seqid5 : ..................................................... : 2597
```

```
                  *      2620          *      2640          *
Seqid1 : AACGCAGCGTTATTACAGTTCGTTCGACCCGAAAGACAAGGACGAAGAAG : 2650
Seqid3 : ..................................................... : 2650
Seqid5 : ..................C..........................C. : 2647

               2660          *      2680          *      2700
Seqid1 : TAACGTGTAATGCTGATAAAACGTTGTGCAACGGCAAATACGGCGGCACA : 2700
Seqid3 : ..................................................... : 2700
Seqid5 : ..................................................... : 2697

                  *      2720          *      2740          *
Seqid1 : AGCAAAAGCGTATTGACCAATTTTGCACGCGGACGCACCTTTTTGATAAC : 2750
Seqid3 : ..................................................... : 2750
Seqid5 : ..................................................... : 2747

               2760
Seqid1 : GATGAGCTACAAGTTTTAA : 2769
Seqid3 : ................... : 2769
Seqid5 : ................... : 2766
```

Figure 2 : Alignment of the BASB024 polypeptide sequences.
Identity to SeqID No:2 is indicated by a dot, and a dash ("-") indicates a missing amino acid.

```
                  *         20          *         40          *
Seqid2 : MRSSFRLKPICFYLMGVMLYHHSYAEDAGRAGSEAQIQVLEDVHVKAKRV :  50
Seqid4 : ..................................................... :  50
Seqid6 : .................T...Y..............................  :  50

                 60          *         80          *        100
Seqid2 : PKDKKVFTDARAVSTRQDIFKSSENLDNIVRSIPGAFTQQDKSSGIVSLN : 100
Seqid4 : ..................................................... : 100
Seqid6 : ..................................................... : 100

                  *        120          *        140          *
Seqid2 : IRGDSGFGRVNTMVDGITQTFYSTSTDAGRAGGSSQFGASVDSNFIAGLD : 150
Seqid4 : ..................................................... : 150
Seqid6 : ..................................................... : 150

                160          *        180          *        200
Seqid2 : VVKGSFSGSAGINSLAGSANLRTLGVDDVVQGNNTYGLLLKGLTGTNSTK : 200
Seqid4 : ..................................................... : 200
Seqid6 : ..................................................... : 200

                  *        220          *        240          *
Seqid2 : GNAMAAIGARKWLESGASVGVLYGHSRRTWAQNYRVGGGGQHIGNFGAEY : 250
Seqid4 : ............................SV..................... : 250
Seqid6 : ............................SV..................... : 250
```

```
                260         *         280         *         300
Seqid2 : LERRKQRYFVQEGGLKFNSNSGKWERDFQRPYWKTKWYQKYNDPQELQKY : 300
Seqid4 : ..................................................... : 300
Seqid6 : .............A.....D.......L..QQ..Y.P.KN..N-...... : 299

                    *       320         *       340         *
Seqid2 : IEGHDKSWRENLAPQYDITPIDPSSLKQQSAGNLFKLEYDGVFNKYTAQF : 350
Seqid4 : ..................................................... : 350
Seqid6 : ..................................................... : 349

                360         *         380         *         400
Seqid2 : RDLNTKIGSRKIINRNYQFNYGLSLNSYANLNLTAAYNSGRQKYPKGSKF : 400
Seqid4 : ..................................................... : 400
Seqid6 : ..........................P.T..................... : 399

                    *       420         *       440         *
Seqid2 : TGWGLLKDFETYNNAKILDLNNTATFRLPRETELQTTLGFNYFHNEYGKN : 450
Seqid4 : ..................................................... : 450
Seqid6 : ..................................................... : 449

                460         *         480         *         500
Seqid2 : RFPEELGLFFDGPDQDNGLYSYLGRFKGDKGLLPQKSTIVQPAGSQYFNT : 500
Seqid4 : ..................................................... : 500
Seqid6 : ..................................................... : 499

                    *       520         *       540         *
Seqid2 : FYFDAALKKDIYRLNYSTNTVGYRFGGEYTGYYGSDDEFKRAFGENSPTY : 550
Seqid4 : ..................................................... : 550
Seqid6 : ..................................................... : 549
```

```
                560         *         580         *         600
Seqid2 : KKHCNQSCGIYEPVLKKYGKKRANNHSVSISADFGDYFMPFASYSRTHRM : 600
Seqid4 : ..................................................... : 600
Seqid6 : .....R............................................... : 599

                    *         620         *         640         *
Seqid2 : PNIQEMYFSQIGDSGVHTALKPERANTWQFGFNTYKKGLLKQDDTLGLKL : 650
Seqid4 : ..................................................... : 650
Seqid6 : ..................................................... : 649

                660         *         680         *         700
Seqid2 : VGYRSRIDNYIHNVYGKWWDLNGNIPSWVSSTGLAYTIQHRNFKDKVHKH : 700
Seqid4 : ..................................................... : 700
Seqid6 : .....................D....................... : 699

                    *         720         *         740         *
Seqid2 : GFELELNYDYGRFFTNLSYAYQKSTQPTNFSDASESPNNASKEDQLKQGY : 750
Seqid4 : ..................................................... : 750
Seqid6 : ..................................................... : 749

                760         *         780         *         800
Seqid2 : GLSRVSALPRDYGRLEVGTRWLGNKLTLGGAMRYFGKSIRATAEERYIDG : 800
Seqid4 : ..................................................... : 800
Seqid6 : ..................................................... : 799

                    *         820         *         840         *
Seqid2 : TNGGNTSNVRQLGKRSIKQTETLARQPLIFDFYAAYEPKKNLIFRAEVKN : 850
Seqid4 : ..................................................... : 850
Seqid6 : ........F............................................ : 849
```

```
                   860         *        880         *        900
Seqid2 : LFDRRYIDPLDAGNDAATQRYYSSFDPKDKDEEVTCNADKTLCNGKYGGT : 900
Seqid4 : .................................................. : 900
Seqid6 : ................................D................. : 899


                *        920
Seqid2 : SKSVLTNFARGRTFLITMSYKF : 922
Seqid4 : ...................... : 922
Seqid6 : ...................... : 921
```

Figure 3. Expression and purification of recombinant BASB024 in *E. coli*.

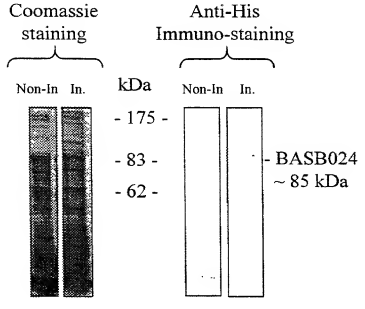
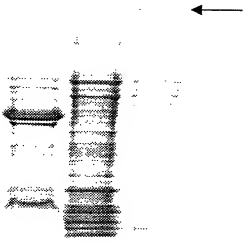


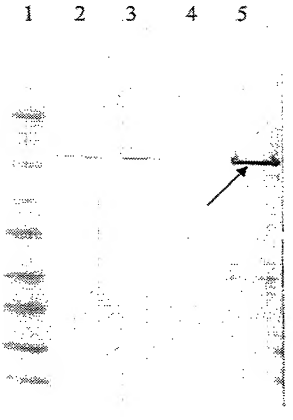
Figure 4 : Coomassie stained SDS-PAGE of the purification fractions of BASB024

1 2 3 4 5



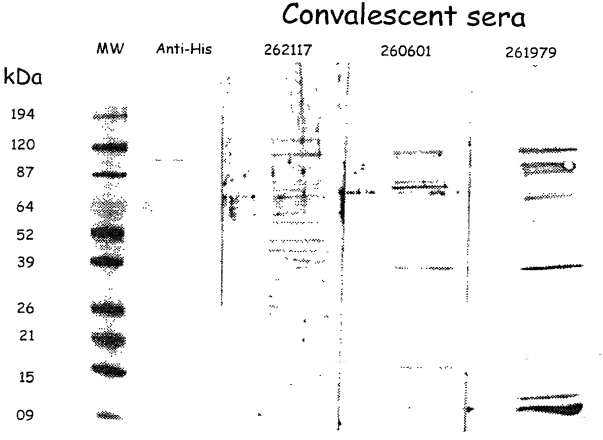
Lanes : 1 : MW : 175, 83, 62, 47.5, 32.5, 25, 16.5, 6.5
2 : Start
3 : Flowthrough
4 : 5 mM imidazole pool
5 : BASB024 enriched pool

Figure 5 : Western blot of purified recombinant BASB024 protein probed with anti-His antibody



Lanes : 1 : MW : 175, 83, 62, 47.5, 32.5, 25, 16.5, 6.5
2 : Start
3 : Flowthrough
4 : 5 mM imidazole pool
5 : BASB024 enriched pool

Figure 6 : Anti-BASB024 antibodies in human convalescent sera by western-blotting using native BASB024 into the gel.



Docket No.: BM45330 PCT/EP99/05989

DECLARATION AND POWER OF ATTORNEY

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name.

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled:

BASB024 Outer Membrane Protein of Neisseria Meningitidis

the specification of which (check one)
[ ] is attached hereto.
[X] was filed on 13 August 1999 as Serial No. PCT/EP99/05989
and was amended on (if applicable).

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to the patentability as defined in Title 37, Code of Federal Regulations, Section 1.56.

I hereby claim foreign priority benefits under Title 35, United States Code, Section 119(a)-(d) or Section 365(b) of any foreign application(s) for patent or inventor's certificate, or Section 365(a) of any PCT International application which designated at least one country other than the United States, listed below and have also identified below any foreign application for patent or Inventor's certificate, or PCT International application having a filing date before that of the application on which priority is claimed.

Prior Foreign Application(s)

Number	Country	Filing Date	Priority Claimed
9818004.5	Great Britain	18 August 1998	Yes

I hereby claim the benefit under Title 35, United States Code, Section 119(e) of any United States provisional application(s) listed below.

Application Number	Filing Date

I hereby claim the benefit under Title 35, United States Code, Section 120 of any United States application(s) or Section 365(c) of any PCT International application

designating the United States, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT International application in the manner provided by the first paragraph of Title 35, United States Code, Section 112, I acknowledge the duty to disclose information which is material to patentability as defined in Title 37, Code of Federal Regulations, Section 1.56 which became available between the filing date of the prior application and the national or PCT international filing date of this application.

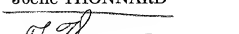
Serial No.	Filing Date	Status

I hereby appoint the practitioners associated with the Customer Number provided below to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith, and direct that all correspondence be addressed to that Customer Number:

Customer Number 25,308

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Full Name of Inventor: Joelle THONNARD

Inventor's Signature: 

Date: 22 February 2001

Residence: Gembloux, Belgium

Citizenship: Belgian

Post Office Address: SmithKline Beecham Corporation
Corporate Intellectual Property - UW2220
P.O. Box 1539
King of Prussia, Pennsylvania 19406-0939

09/762926
528 Rec'd PCT/PTO 14 FEB 2001
PCT/EP99/05989

SEQUENCE LISTING

<110> SmithKline Beecham Biologicals S.A.

<120> Novel compounds

<130> BM45330

<160> 9

<170> FastSEQ for Windows Version 3.0

<210> 1
<211> 2769
<212> DNA
<213> Bacteria

<400> 1

atgagatctt ctttccggtt gaagccgatt tgtttttatc ttatgggtgt tatgctatat 60
catcatagtt atgccgaaga tgcagggcgc gcgggcagcg aggcgcagat acaggttttg 120
gaagatgtgc acgtcaaggc gaagcgcgta ccgaaagaca aaaaagtgtt taccgatgcg 180
cgtgccgtat cgacccgtca ggatatattc aaatccagcg aaaacctcga caacatcgta 240
cgcagcatcc ccggtgcgtt tacacagcaa gataaaagct cgggcattgt gtctttgaat 300
attcgcggcg acagcgggtt cgggcgggtc aatacgatgg tggacggcat cacgcagacc 360
ttttattcga cttctaccga tgcgggcagg gcaggcggtt catctcaatt cggtgcatct 420
gtcgacagca attttattgc cggactggat gtcgtcaaag gcagctttcag cggctcggca 480
ggcatcaaca gccttgccgg ttcggcgaat ctgcggactt taggcgtgga tgacgtcgtt 540
cagggcaata atacctacgg cctgctgcta aaaggtctga ccggcaccaa ttcaaccaaa 600
ggtaatgcga tggcggcgat aggtgcgcgc aaatggctgg aaagcggagc atctgtcggt 660
gtgctttacg ggcacagcag gcgcacgtgg gcgcaaaatt accgcgtggg cggcggcggg 720
cagcacatcg gaaattttgg cgcggaatat ctggaacggc gcaaacagcg atattttgta 780
caagaaggcg ggttgaaatt caattccaac agcggaaaat gggagcggga tttccaaagg 840
ccgtactgga aaaccaagtg gtatcaaaaa tacaatgacc cccaagaact gcaaaaatac 900
atcgaaggtc atgacaaaag ctggcgggaa aacctggcgc cgcaatacga catcacccca 960
atcgatccgt ccagcctgaa gcagcagtcg gcaggcaatc tgtttaaatt ggaatacgac 1020

```
ggcgtattca ataaatacac ggcgcaattt cgcgatttaa acaccaaaat cggcagccgc   1080
aaaatcatca accgcaatta tcaattcaat tacggtttat ctttaaactc atatgccaac   1140
ctcaatctga ccgcagccta caattcgggc aggcagaaat atccgaaagg gtcgaagttt   1200
acaggctggg ggcttttaaa agattttgaa acctacaaca acgcgaaaat cctcgacctc   1260
aacaacaccg ccacctttcg gctgccccgc gaaaccgagt tgcaaaccac tttgggcttc   1320
aattatttcc acaacgaata cggcaaaaac cgctttcctg aagaattggg gctgtttttc   1380
gacggtccgg atcaggacaa cgggctttat tcctatttgg ggcggtttaa gggcgataaa   1440
gggctgctgc cccaaaaatc aaccatcgtc caaccggccg gcagccaata tttcaacacg   1500
ttctacttcg atgccgcgct caaaaaagac atttaccgct taaactacag caccaatacc   1560
gtcggctacc gtttcggcgg cgaatatacg ggctattacg gctcggatga cgaatttaag   1620
cgggcattcg gagaaaactc gccgacatac aagaaacatt gcaaccagag ctgcggaatt   1680
tatgaacccg tattgaaaaa atacggcaaa aagcgcgcca acaaccattc ggtcagcatt   1740
agtgcggact tcggcgatta tttcatgccg ttcgccagct attcgcgcac acaccgtatg   1800
cccaacatcc aagaaatgta tttttcccaa atcggcgact ccggcgttca caccgcctta   1860
aaaccagagc gcgcaaacac ttggcaattt ggcttcaata cctataaaaa aggattgtta   1920
aaacaagatg atacattagg attaaaactg gtcggctacc gcagccgcat cgacaactac   1980
atccacaacg tttacgggaa atggtgggat ttgaacggga atattccgag ctgggtcagc   2040
agcaccgggc ttgcctacac catccaacac cgcaatttca aagacaaagt acacaaacac   2100
ggttttgagt tggagctgaa ttacgattat gggcgttttt tcaccaacct ttcttacgcc   2160
tatcaaaaaa gcacgcaacc gaccaacttc agcgatgcga gcgaatcgcc caacaatgcg   2220
tccaaagaag accaactcaa acaaggttat gggttgagca gggtttccgc cctgccgcga   2280
gattacggac gtttggaagt cggtacgcgc tggttgggca acaaactgac tttgggcggc   2340
gcgatgcgct atttcggcaa gagcatccgc gcgacggctg aagaacgcta tatcgacggc   2400
accaacgggg gaaataccag caatgtccgg caactgggca agcgttccat caaacaaacc   2460
gaaacccttg cccgccagcc tttgattttt gatttttacg ccgcttacga gccgaagaaa   2520
aaccttattt tccgcgccga agtcaaaaat ctgttcgaca ggcgttatat cgatccgctc   2580
gatgcgggca atgatgcggc aacgcagcgt tattacagtt cgttcgaccc gaaagacaag   2640
gacgaagaag taacgtgtaa tgctgataaa acgttgtgca acggcaaata cggcggcaca   2700
agcaaaagcg tattgaccaa ttttgcacgc ggacgcacct ttttgataac gatgagctac   2760
aagttttaa                                                          2769
```

<210> 2
<211> 922
<212> PRT
<213> Bacteria

<400> 2

```
Met Arg Ser Ser Phe Arg Leu Lys Pro Ile Cys Phe Tyr Leu Met Gly
 1               5                  10                  15
Val Met Leu Tyr His His Ser Tyr Ala Glu Asp Ala Gly Arg Ala Gly
            20                  25                  30
```

```
Ser Glu Ala Gln Ile Gln Val Leu Glu Asp Val His Val Lys Ala Lys
        35                  40                  45
Arg Val Pro Lys Asp Lys Lys Val Phe Thr Asp Ala Arg Ala Val Ser
    50                  55                  60
Thr Arg Gln Asp Ile Phe Lys Ser Ser Glu Asn Leu Asp Asn Ile Val
65                  70                  75                  80
Arg Ser Ile Pro Gly Ala Phe Thr Gln Gln Asp Lys Ser Ser Gly Ile
                85                  90                  95
Val Ser Leu Asn Ile Arg Gly Asp Ser Gly Phe Gly Arg Val Asn Thr
            100                 105                 110
Met Val Asp Gly Ile Thr Gln Thr Phe Tyr Ser Thr Ser Thr Asp Ala
        115                 120                 125
Gly Arg Ala Gly Gly Ser Ser Gln Phe Gly Ala Ser Val Asp Ser Asn
    130                 135                 140
Phe Ile Ala Gly Leu Asp Val Val Lys Gly Ser Phe Ser Gly Ser Ala
145                 150                 155                 160
Gly Ile Asn Ser Leu Ala Gly Ser Ala Asn Leu Arg Thr Leu Gly Val
                165                 170                 175
Asp Asp Val Val Gln Gly Asn Asn Thr Tyr Gly Leu Leu Leu Lys Gly
            180                 185                 190
Leu Thr Gly Thr Asn Ser Thr Lys Gly Asn Ala Met Ala Ala Ile Gly
        195                 200                 205
Ala Arg Lys Trp Leu Glu Ser Gly Ala Ser Val Gly Val Leu Tyr Gly
    210                 215                 220
His Ser Arg Arg Thr Trp Ala Gln Asn Tyr Arg Val Gly Gly Gly Gly
225                 230                 235                 240
Gln His Ile Gly Asn Phe Gly Ala Glu Tyr Leu Glu Arg Arg Lys Gln
                245                 250                 255
Arg Tyr Phe Val Gln Glu Gly Gly Leu Lys Phe Asn Ser Asn Ser Gly
            260                 265                 270
Lys Trp Glu Arg Asp Phe Gln Arg Pro Tyr Trp Lys Thr Lys Trp Tyr
        275                 280                 285
Gln Lys Tyr Asn Asp Pro Gln Glu Leu Gln Lys Tyr Ile Glu Gly His
    290                 295                 300
Asp Lys Ser Trp Arg Glu Asn Leu Ala Pro Gln Tyr Asp Ile Thr Pro
305                 310                 315                 320
Ile Asp Pro Ser Ser Leu Lys Gln Gln Ser Ala Gly Asn Leu Phe Lys
                325                 330                 335
Leu Glu Tyr Asp Gly Val Phe Asn Lys Tyr Thr Ala Gln Phe Arg Asp
            340                 345                 350
Leu Asn Thr Lys Ile Gly Ser Arg Lys Ile Ile Asn Arg Asn Tyr Gln
```

```
        355                 360                 365
Phe Asn Tyr Gly Leu Ser Leu Asn Ser Tyr Ala Asn Leu Asn Leu Thr
    370                 375                 380
Ala Ala Tyr Asn Ser Gly Arg Gln Lys Tyr Pro Lys Gly Ser Lys Phe
385                 390                 395                 400
Thr Gly Trp Gly Leu Leu Lys Asp Phe Glu Thr Tyr Asn Asn Ala Lys
                405                 410                 415
Ile Leu Asp Leu Asn Asn Thr Ala Thr Phe Arg Leu Pro Arg Glu Thr
            420                 425                 430
Glu Leu Gln Thr Thr Leu Gly Phe Asn Tyr Phe His Asn Glu Tyr Gly
        435                 440                 445
Lys Asn Arg Phe Pro Glu Glu Leu Gly Leu Phe Phe Asp Gly Pro Asp
    450                 455                 460
Gln Asp Asn Gly Leu Tyr Ser Tyr Leu Gly Arg Phe Lys Gly Asp Lys
465                 470                 475                 480
Gly Leu Leu Pro Gln Lys Ser Thr Ile Val Gln Pro Ala Gly Ser Gln
                485                 490                 495
Tyr Phe Asn Thr Phe Tyr Phe Asp Ala Ala Leu Lys Lys Asp Ile Tyr
            500                 505                 510
Arg Leu Asn Tyr Ser Thr Asn Thr Val Gly Tyr Arg Phe Gly Gly Glu
        515                 520                 525
Tyr Thr Gly Tyr Tyr Gly Ser Asp Asp Glu Phe Lys Arg Ala Phe Gly
    530                 535                 540
Glu Asn Ser Pro Thr Tyr Lys Lys His Cys Asn Gln Ser Cys Gly Ile
545                 550                 555                 560
Tyr Glu Pro Val Leu Lys Lys Tyr Gly Lys Lys Arg Ala Asn Asn His
                565                 570                 575
Ser Val Ser Ile Ser Ala Asp Phe Gly Asp Tyr Phe Met Pro Phe Ala
            580                 585                 590
Ser Tyr Ser Arg Thr His Arg Met Pro Asn Ile Gln Glu Met Tyr Phe
        595                 600                 605
Ser Gln Ile Gly Asp Ser Gly Val His Thr Ala Leu Lys Pro Glu Arg
    610                 615                 620
Ala Asn Thr Trp Gln Phe Gly Phe Asn Thr Tyr Lys Lys Gly Leu Leu
625                 630                 635                 640
Lys Gln Asp Asp Thr Leu Gly Leu Lys Leu Val Gly Tyr Arg Ser Arg
                645                 650                 655
Ile Asp Asn Tyr Ile His Asn Val Tyr Gly Lys Trp Trp Asp Leu Asn
            660                 665                 670
Gly Asn Ile Pro Ser Trp Val Ser Ser Thr Gly Leu Ala Tyr Thr Ile
        675                 680                 685
```

```
Gln His Arg Asn Phe Lys Asp Lys Val His Lys His Gly Phe Glu Leu
    690                 695                 700
Glu Leu Asn Tyr Asp Tyr Gly Arg Phe Phe Thr Asn Leu Ser Tyr Ala
705                 710                 715                 720
Tyr Gln Lys Ser Thr Gln Pro Thr Asn Phe Ser Asp Ala Ser Glu Ser
                725                 730                 735
Pro Asn Asn Ala Ser Lys Glu Asp Gln Leu Lys Gln Gly Tyr Gly Leu
            740                 745                 750
Ser Arg Val Ser Ala Leu Pro Arg Asp Tyr Gly Arg Leu Glu Val Gly
        755                 760                 765
Thr Arg Trp Leu Gly Asn Lys Leu Thr Leu Gly Gly Ala Met Arg Tyr
    770                 775                 780
Phe Gly Lys Ser Ile Arg Ala Thr Ala Glu Glu Arg Tyr Ile Asp Gly
785                 790                 795                 800
Thr Asn Gly Gly Asn Thr Ser Asn Val Arg Gln Leu Gly Lys Arg Ser
                805                 810                 815
Ile Lys Gln Thr Glu Thr Leu Ala Arg Gln Pro Leu Ile Phe Asp Phe
            820                 825                 830
Tyr Ala Ala Tyr Glu Pro Lys Lys Asn Leu Ile Phe Arg Ala Glu Val
        835                 840                 845
Lys Asn Leu Phe Asp Arg Arg Tyr Ile Asp Pro Leu Asp Ala Gly Asn
    850                 855                 860
Asp Ala Ala Thr Gln Arg Tyr Tyr Ser Ser Phe Asp Pro Lys Asp Lys
865                 870                 875                 880
Asp Glu Glu Val Thr Cys Asn Ala Asp Lys Thr Leu Cys Asn Gly Lys
                885                 890                 895
Tyr Gly Gly Thr Ser Lys Ser Val Leu Thr Asn Phe Ala Arg Gly Arg
            900                 905                 910
Thr Phe Leu Ile Thr Met Ser Tyr Lys Phe
        915                 920
```

<210> 3
<211> 2769
<212> DNA
<213> Bacteria

<400> 3

```
atgagatctt ctttccggtt gaagccgatt tgtttttatc ttatgggtgt tatgctatat      60
catcatagtt atgccgaaga tgcagggcgc gcgggcagcg aggcgcagat acaggttttg     120
gaagatgtgc acgtcaaggc gaagcgcgta ccgaaagaca aaaaagtgtt taccgatgcg     180
cgtgccgtat cgacccgtca ggatatattc aaatccagcg aaaacctcga caacatcgta     240
```

```
cgcagcatcc ccggtgcgtt tacacagcaa gataaaagct cgggcattgt gtctttgaat    300
attcgcggcg acagcgggtt cgggcgggtc aatacgatgg tggacggcat cacgcagacc    360
ttttattcga cttctaccga tgcgggcagg gcaggcggtt catctcaatt cggtgcatct    420
gtcgacagca attttattgc cggactggat gtcgtcaaag gcagcttcag cggctcggca    480
ggcatcaaca gccttgccgg ttcggcgaat ctgcggactt taggcgtgga tgacgtcgtt    540
cagggcaata atacctacgg cctgctgcta aaaggtctga ccggcaccaa ttcaaccaaa    600
ggtaatgcga tggcggcgat aggtgcgcgc aaatggctgg aaagcggagc atctgtcggt    660
gtgctttacg ggcacagcag gcgcagcgtg gcgcaaaatt accgcgtggg cggcggcggg    720
cagcacatcg gaaattttgg cgcggaatat ctggaacggc gcaagcagcg atattttgta    780
caagaaggcg ggttgaaatt caattccaac agcggaaaat gggagcggga tttccaaagg    840
ccgtactgga aaaccaagtg gtatcaaaaa tacaatgacc cccaagaact gcaaaaatac    900
atcgaaggtc atgacaaaag ctggcgggaa aacctggcgc cgcaatacga catcaccccc    960
atcgatccgt ccagcctgaa gcagcagtcg gcaggcaatc tgtttaaatt ggaatacgac   1020
ggcgtattca ataaatacac ggcgcaattt cgcgatttaa acaccaaaat cggcagccgc   1080
aaaatcatca accgcaatta tcaattcaat tacggtttat ctttaaactc atatgccaac   1140
ctcaatctga ccgcagccta caattcgggc aggcagaaat atccgaaagg gtcgaagttt   1200
acaggctggg ggcttttaaa agattttgaa acctacaaca acgcgaaaat cctcgacctc   1260
aacaacaccg ccacctccg gctgccccgc gaaaccgagt tgcaaaccac tttgggcttc   1320
aattatttcc acaacgaata cggcaaaaac cgctttcctg aagaattggg gctgtttttc   1380
gacggtccgg atcaggacaa cgggctttat tcctatttgg ggcggtttaa gggcgataaa   1440
gggctgctgc cccaaaaatc aaccatcgtc caaccggccg gcagccaata tttcaacacg   1500
ttctacttcg atgccgcgct caaaaaagac atttaccgct taaactacag caccaatacc   1560
gtcggctacc gtttcggcgg cgaatatacg ggctattacg gctcggatga cgaatttaag   1620
cgggcattcg gagaaaactc gccgacatac aagaaacatt gcaaccagag ctgcggaatt   1680
tatgaacccg tattgaaaaa atacggcaaa aagcgcgcca acaaccattc ggtcagcatt   1740
agtgcggact tcggcgatta tttcatgccg ttcgccagct attcgcgcac acaccgtatg   1800
cccaacatcc aagaaatgta ttttttcccaa atcggcgact ccggcgttca caccgcctta   1860
aaaccagagc gcgcaaacac ttggcaattt ggcttcaata cctataaaaa aggattgtta   1920
aaacaagatg atacattagg attaaaactg gtcggctacc gcagccgcat cgacaactac   1980
atccacaacg tttacgggaa atggtgggat ttgaacggga atattccgag ctgggtcagc   2040
agcaccgggc ttgcctacac catccaacac cgcaatttca aagacaaagt acacaaacac   2100
ggttttgagt tggagctgaa ttacgattat gggcgttttt tcaccaacct ttcttacgcc   2160
tatcaaaaaa gcacgcaacc gaccaacttc agcgatgcga gcgaatcgcc caacaatgcg   2220
tccaaagaag accaactcaa acaaggttat gggttgagca gggttccgc cctgccgcga   2280
gattacggac gtttggaagt cggtacgcgc tggttgggca acaaactgac tttgggcggc   2340
gcgatgcgct atttcggcaa gagcatccgc gcgacggctg aagaacgcta tatcgacggc   2400
accaacgggg gaaataccag caatgtccgg caactgggca agcgttccat caaacaaacc   2460
gaaacccttg cccgccagcc tttgattttt gatttttacg ccgcttacga gccgaagaaa   2520
aaccttattt tccgcgccga agtcaaaaat ctgttcgaca ggcgttatat cgatccgctc   2580
gatgcgggca atgatgcggc aacgcagcgt tattacagtt cgttcgaccc gaaagacaag   2640
gacgaagaag taacgtgtaa tgctgataaa acgttgtgca acggcaaata cggcggcaca   2700
```

agcaaaagcg tattgaccaa ttttgcacgc ggacgcacct ttttgataac gatgagctac 2760
aagttttaa 2769

<210> 4
<211> 922
<212> PRT
<213> Bacteria

<400> 4

Met Arg Ser Ser Phe Arg Leu Lys Pro Ile Cys Phe Tyr Leu Met Gly
1 5 10 15
Val Met Leu Tyr His His Ser Tyr Ala Glu Asp Ala Gly Arg Ala Gly
20 25 30
Ser Glu Ala Gln Ile Gln Val Leu Glu Asp Val His Val Lys Ala Lys
35 40 45
Arg Val Pro Lys Asp Lys Lys Val Phe Thr Asp Ala Arg Ala Val Ser
50 55 60
Thr Arg Gln Asp Ile Phe Lys Ser Ser Glu Asn Leu Asp Asn Ile Val
65 70 75 80
Arg Ser Ile Pro Gly Ala Phe Thr Gln Gln Asp Lys Ser Ser Gly Ile
85 90 95
Val Ser Leu Asn Ile Arg Gly Asp Ser Gly Phe Gly Arg Val Asn Thr
100 105 110
Met Val Asp Gly Ile Thr Gln Thr Phe Tyr Ser Thr Ser Thr Asp Ala
115 120 125
Gly Arg Ala Gly Gly Ser Ser Gln Phe Gly Ala Ser Val Asp Ser Asn
130 135 140
Phe Ile Ala Gly Leu Asp Val Val Lys Gly Ser Phe Ser Gly Ser Ala
145 150 155 160
Gly Ile Asn Ser Leu Ala Gly Ser Ala Asn Leu Arg Thr Leu Gly Val
165 170 175
Asp Asp Val Val Gln Gly Asn Asn Thr Tyr Gly Leu Leu Leu Lys Gly
180 185 190
Leu Thr Gly Thr Asn Ser Thr Lys Gly Asn Ala Met Ala Ala Ile Gly
195 200 205
Ala Arg Lys Trp Leu Glu Ser Gly Ala Ser Val Gly Val Leu Tyr Gly
210 215 220
His Ser Arg Arg Ser Val Ala Gln Asn Tyr Arg Val Gly Gly Gly Gly
225 230 235 240
Gln His Ile Gly Asn Phe Gly Ala Glu Tyr Leu Glu Arg Arg Lys Gln
245 250 255

```
Arg Tyr Phe Val Gln Glu Gly Gly Leu Lys Phe Asn Ser Asn Ser Gly
            260                 265                 270
Lys Trp Glu Arg Asp Phe Gln Arg Pro Tyr Trp Lys Thr Lys Trp Tyr
        275                 280                 285
Gln Lys Tyr Asn Asp Pro Gln Glu Leu Gln Lys Tyr Ile Glu Gly His
    290                 295                 300
Asp Lys Ser Trp Arg Glu Asn Leu Ala Pro Gln Tyr Asp Ile Thr Pro
305                 310                 315                 320
Ile Asp Pro Ser Ser Leu Lys Gln Gln Ser Ala Gly Asn Leu Phe Lys
                325                 330                 335
Leu Glu Tyr Asp Gly Val Phe Asn Lys Tyr Thr Ala Gln Phe Arg Asp
            340                 345                 350
Leu Asn Thr Lys Ile Gly Ser Arg Lys Ile Ile Asn Arg Asn Tyr Gln
        355                 360                 365
Phe Asn Tyr Gly Leu Ser Leu Asn Ser Tyr Ala Asn Leu Asn Leu Thr
    370                 375                 380
Ala Ala Tyr Asn Ser Gly Arg Gln Lys Tyr Pro Lys Gly Ser Lys Phe
385                 390                 395                 400
Thr Gly Trp Gly Leu Leu Lys Asp Phe Glu Thr Tyr Asn Asn Ala Lys
                405                 410                 415
Ile Leu Asp Leu Asn Asn Thr Ala Thr Phe Arg Leu Pro Arg Glu Thr
            420                 425                 430
Glu Leu Gln Thr Thr Leu Gly Phe Asn Tyr Phe His Asn Glu Tyr Gly
        435                 440                 445
Lys Asn Arg Phe Pro Glu Glu Leu Gly Leu Phe Phe Asp Gly Pro Asp
    450                 455                 460
Gln Asp Asn Gly Leu Tyr Ser Tyr Leu Gly Arg Phe Lys Gly Asp Lys
465                 470                 475                 480
Gly Leu Leu Pro Gln Lys Ser Thr Ile Val Gln Pro Ala Gly Ser Gln
                485                 490                 495
Tyr Phe Asn Thr Phe Tyr Phe Asp Ala Ala Leu Lys Lys Asp Ile Tyr
            500                 505                 510
Arg Leu Asn Tyr Ser Thr Asn Thr Val Gly Tyr Arg Phe Gly Gly Glu
        515                 520                 525
Tyr Thr Gly Tyr Tyr Gly Ser Asp Asp Glu Phe Lys Arg Ala Phe Gly
    530                 535                 540
Glu Asn Ser Pro Thr Tyr Lys Lys His Cys Asn Gln Ser Cys Gly Ile
545                 550                 555                 560
Tyr Glu Pro Val Leu Lys Lys Tyr Gly Lys Lys Arg Ala Asn Asn His
                565                 570                 575
Ser Val Ser Ile Ser Ala Asp Phe Gly Asp Tyr Phe Met Pro Phe Ala
```

```
            580                 585                 590
Ser Tyr Ser Arg Thr His Arg Met Pro Asn Ile Gln Glu Met Tyr Phe
        595                 600                 605
Ser Gln Ile Gly Asp Ser Gly Val His Thr Ala Leu Lys Pro Glu Arg
    610                 615                 620
Ala Asn Thr Trp Gln Phe Gly Phe Asn Thr Tyr Lys Lys Gly Leu Leu
625                 630                 635                 640
Lys Gln Asp Asp Thr Leu Gly Leu Lys Leu Val Gly Tyr Arg Ser Arg
                645                 650                 655
Ile Asp Asn Tyr Ile His Asn Val Tyr Gly Lys Trp Trp Asp Leu Asn
            660                 665                 670
Gly Asn Ile Pro Ser Trp Val Ser Ser Thr Gly Leu Ala Tyr Thr Ile
        675                 680                 685
Gln His Arg Asn Phe Lys Asp Lys Val His Lys His Gly Phe Glu Leu
    690                 695                 700
Glu Leu Asn Tyr Asp Tyr Gly Arg Phe Phe Thr Asn Leu Ser Tyr Ala
705                 710                 715                 720
Tyr Gln Lys Ser Thr Gln Pro Thr Asn Phe Ser Asp Ala Ser Glu Ser
                725                 730                 735
Pro Asn Asn Ala Ser Lys Glu Asp Gln Leu Lys Gln Gly Tyr Gly Leu
            740                 745                 750
Ser Arg Val Ser Ala Leu Pro Arg Asp Tyr Gly Arg Leu Glu Val Gly
        755                 760                 765
Thr Arg Trp Leu Gly Asn Lys Leu Thr Leu Gly Gly Ala Met Arg Tyr
    770                 775                 780
Phe Gly Lys Ser Ile Arg Ala Thr Ala Glu Glu Arg Tyr Ile Asp Gly
785                 790                 795                 800
Thr Asn Gly Gly Asn Thr Ser Asn Val Arg Gln Leu Gly Lys Arg Ser
                805                 810                 815
Ile Lys Gln Thr Glu Thr Leu Ala Arg Gln Pro Leu Ile Phe Asp Phe
            820                 825                 830
Tyr Ala Ala Tyr Glu Pro Lys Lys Asn Leu Ile Phe Arg Ala Glu Val
        835                 840                 845
Lys Asn Leu Phe Asp Arg Arg Tyr Ile Asp Pro Leu Asp Ala Gly Asn
    850                 855                 860
Asp Ala Ala Thr Gln Arg Tyr Tyr Ser Ser Phe Asp Pro Lys Asp Lys
865                 870                 875                 880
Asp Glu Glu Val Thr Cys Asn Ala Asp Lys Thr Leu Cys Asn Gly Lys
                885                 890                 895
Tyr Gly Gly Thr Ser Lys Ser Val Leu Thr Asn Phe Ala Arg Gly Arg
            900                 905                 910
```

Thr Phe Leu Ile Thr Met Ser Tyr Lys Phe
915 920

<210> 5
<211> 2766
<212> DNA
<213> Bacteria

<400> 5

atgagatctt ctttccggtt gaagccgatt tgtttttacc ttatgggtgt tacgctatat 60
cattatagtt atgccgaaga tgcagggcgc gcgggcagcg aggcgcagat acaggttttg 120
gaagatgtgc acgtcaaggc gaagcgcgta ccgaaagaca aaaaagtgtt taccgatgcg 180
cgtgccgtat cgacccgtca ggatatattc aaatccagcg aaaacctcga caacatcgta 240
cgcagcatcc ccggtgcgtt tacacagcaa gataaaagct cgggcattgt gtctttgaat 300
attcgcggcg acagcgggtt cgggcgggtc aatacgatgg tggacggcat cacgcagacc 360
ttttattcga cttctaccga tgcgggcagg gcaggcggtt catctcaatt cggtgcatct 420
gtcgacagca attttattgc cggactggat gtcgtcaaag gcagcttcag cggctcggca 480
ggcatcaaca gccttgccgg ttcggcgaat ctgcggactt taggcgtgga tgacgtcgtt 540
cagggcaata atacctacgg cctgctgcta aaaggtctga ccggcaccaa ttcaaccaaa 600
ggtaatgcga tggcggcgat aggtgcgcgc aaatggctgg aaagcggagc atctgtcggt 660
gtgctttacg ggcacagcag gcgcagcgtg gcgcaaaatt accgcgtggg cggcggcggg 720
cagcacatcg gaaattttgg cgcggaatat ttggaacggc gcaagcagcg atattttgta 780
caagagggtg ctttgaaatt caatttccgac agcggaaaat gggagcggga tttacaaagg 840
caacagtgga aatacaagcc gtataaaaat tacaacaacc aagaactaca aaaatacatc 900
gaaggtcatg acaaaagctg gcgggaaaac ctggcgccgc aatacgacat caccccccatc 960
gatccgtcca gcctgaagca gcagtcggca ggcaatctgt ttaaattgga atacgacggc 1020
gtattcaata aatacacggc gcaatttcgc gatttaaaca ccaaaatcgg cagccgcaaa 1080
atcatcaacc gcaattatca gttcaattac ggtttgtctt tgaacccgta taccaacctc 1140
aatctgaccg cagcctacaa ttcgggcagg cagaaatatc cgaaagggtc gaagtttaca 1200
ggctgggggc ttttaaagga ttttgaaacc tacaacaacg cgaaaatcct cgacctcaac 1260
aacaccgcca ccttccggct gccccgcgaa accgagttgc aaaccacttt gggcttcaat 1320
tatttccaca acgaatacgg caaaaaccgc tttcctgaag aattggggct gtttttcgac 1380
ggtcctgatc aggacaacgg gctttattcc tatttggggc ggtttaaggg cgataaaggg 1440
ctgctgcccc aaaaatcaac cattgtccaa ccggccggca gccaatattt caacacgttc 1500
tacttcgatg ccgcgctcaa aaaagacatt taccgcttaa actacagcac caatacgtc 1560
ggctaccgtt tcggcggcga atatacgggc tattacggct cggatgacga atttaagcgg 1620
gcattcggag aaaactcgcc gacatacaag aaacattgca accggagctg cgggatttat 1680
gaacccgtat tgaaaaaata cggcaaaaag cgcgccaaca accattcggt cagcattagt 1740
gcggacttcg gcgattattt catgccgttc gccagctatt cgcgcacaca ccgtatgccc 1800
aacatccaag aaatgtattt ttcccaaatc ggcgactccg gcgttcacac cgccttaaaa 1860
ccagagcgcg caaacacttg gcaatttggc ttcaatacct ataaaaaagg attgttaaaa 1920

caagatgata cattaggatt aaaactggtc ggctaccgca gccgcatcga caactacatc 1980
cacaacgttt acgggaaatg gtgggatttg aacggggata ttccgagctg ggtcagcagc 2040
acgggcttg cctacaccat ccaacatcgc aatttcaaag acaaagtgca caaacacggt 2100
tttgagttgg agctgaatta cgattatggg cgtttttttca ccaacctttc ttacgcctat 2160
caaaaaagca cgcaaccgac caacttcagc gatgcgagcg aatcgcccaa caatgcgtcc 2220
aaagaagacc aactcaaaca aggttatggg ttgagcaggg tttccgccct gccgcgagat 2280
tacggacgtt tggaagtcgg tacgcgctgg ttgggcaaca aactgacttt gggcggcgcg 2340
atgcgctatt tcggcaagag catccgcgcg acggctgaag aacgctatat cgacggcacc 2400
aacgggggaa ataccagcaa tttccggcaa ctgggcaagc gttccatcaa acaaaccgaa 2460
actcttgccc gccagccttt gatttttgat ttttacgccg cttacgagcc gaagaaaaac 2520
cttattttcc gcgccgaagt caaaaatctg ttcgacaggc gttatatcga tccgctcgat 2580
gcgggcaatg atgcggcaac gcagcgttat tacagctcgt tcgacccgaa agacaaggac 2640
gaagacgtaa cgtgtaatgc tgataaaacg ttgtgcaacg gcaaatacgg cggcacaagc 2700
aaaagcgtat tgaccaattt tgcacgcgga cgcacctttt tgataacgat gagctacaag 2760
ttttaa 2766

<210> 6
<211> 921
<212> PRT
<213> Bacteria

<400> 6

```
Met Arg Ser Ser Phe Arg Leu Lys Pro Ile Cys Phe Tyr Leu Met Gly
 1               5                  10                  15
Val Thr Leu Tyr His Tyr Ser Tyr Ala Glu Asp Ala Gly Arg Ala Gly
            20                  25                  30
Ser Glu Ala Gln Ile Gln Val Leu Glu Asp Val His Val Lys Ala Lys
        35                  40                  45
Arg Val Pro Lys Asp Lys Lys Val Phe Thr Asp Ala Arg Ala Val Ser
    50                  55                  60
Thr Arg Gln Asp Ile Phe Lys Ser Ser Glu Asn Leu Asp Asn Ile Val
65                  70                  75                  80
Arg Ser Ile Pro Gly Ala Phe Thr Gln Gln Asp Lys Ser Ser Gly Ile
                85                  90                  95
Val Ser Leu Asn Ile Arg Gly Asp Ser Gly Phe Gly Arg Val Asn Thr
            100                 105                 110
Met Val Asp Gly Ile Thr Gln Thr Phe Tyr Ser Thr Ser Thr Asp Ala
        115                 120                 125
Gly Arg Ala Gly Gly Ser Ser Gln Phe Gly Ala Ser Val Asp Ser Asn
    130                 135                 140
Phe Ile Ala Gly Leu Asp Val Val Lys Gly Ser Phe Ser Gly Ser Ala
```

```
145                 150                 155                 160
Gly Ile Asn Ser Leu Ala Gly Ser Ala Asn Leu Arg Thr Leu Gly Val
                165                 170                 175
Asp Asp Val Val Gln Gly Asn Asn Thr Tyr Gly Leu Leu Leu Lys Gly
            180                 185                 190
Leu Thr Gly Thr Asn Ser Thr Lys Gly Asn Ala Met Ala Ala Ile Gly
        195                 200                 205
Ala Arg Lys Trp Leu Glu Ser Gly Ala Ser Val Gly Val Leu Tyr Gly
    210                 215                 220
His Ser Arg Arg Ser Val Ala Gln Asn Tyr Arg Val Gly Gly Gly Gly
225                 230                 235                 240
Gln His Ile Gly Asn Phe Gly Ala Glu Tyr Leu Glu Arg Arg Lys Gln
                245                 250                 255
Arg Tyr Phe Val Gln Glu Gly Ala Leu Lys Phe Asn Ser Asp Ser Gly
            260                 265                 270
Lys Trp Glu Arg Asp Leu Gln Arg Gln Gln Trp Lys Tyr Lys Pro Tyr
        275                 280                 285
Lys Asn Tyr Asn Asn Gln Glu Leu Gln Lys Tyr Ile Glu Gly His Asp
    290                 295                 300
Lys Ser Trp Arg Glu Asn Leu Ala Pro Gln Tyr Asp Ile Thr Pro Ile
305                 310                 315                 320
Asp Pro Ser Ser Leu Lys Gln Gln Ser Ala Gly Asn Leu Phe Lys Leu
                325                 330                 335
Glu Tyr Asp Gly Val Phe Asn Lys Tyr Thr Ala Gln Phe Arg Asp Leu
            340                 345                 350
Asn Thr Lys Ile Gly Ser Arg Lys Ile Ile Asn Arg Asn Tyr Gln Phe
        355                 360                 365
Asn Tyr Gly Leu Ser Leu Asn Pro Tyr Thr Asn Leu Asn Leu Thr Ala
    370                 375                 380
Ala Tyr Asn Ser Gly Arg Gln Lys Tyr Pro Lys Gly Ser Lys Phe Thr
385                 390                 395                 400
Gly Trp Gly Leu Leu Lys Asp Phe Glu Thr Tyr Asn Asn Ala Lys Ile
                405                 410                 415
Leu Asp Leu Asn Asn Thr Ala Thr Phe Arg Leu Pro Arg Glu Thr Glu
            420                 425                 430
Leu Gln Thr Thr Leu Gly Phe Asn Tyr Phe His Asn Glu Tyr Gly Lys
        435                 440                 445
Asn Arg Phe Pro Glu Glu Leu Gly Leu Phe Phe Asp Gly Pro Asp Gln
    450                 455                 460
Asp Asn Gly Leu Tyr Ser Tyr Leu Gly Arg Phe Lys Gly Asp Lys Gly
465                 470                 475                 480
```

```
Leu Leu Pro Gln Lys Ser Thr Ile Val Gln Pro Ala Gly Ser Gln Tyr
                485                 490                 495
Phe Asn Thr Phe Tyr Phe Asp Ala Ala Leu Lys Lys Asp Ile Tyr Arg
            500                 505                 510
Leu Asn Tyr Ser Thr Asn Thr Val Gly Tyr Arg Phe Gly Gly Glu Tyr
        515                 520                 525
Thr Gly Tyr Tyr Gly Ser Asp Asp Glu Phe Lys Arg Ala Phe Gly Glu
    530                 535                 540
Asn Ser Pro Thr Tyr Lys Lys His Cys Asn Arg Ser Cys Gly Ile Tyr
545                 550                 555                 560
Glu Pro Val Leu Lys Lys Tyr Gly Lys Lys Arg Ala Asn Asn His Ser
                565                 570                 575
Val Ser Ile Ser Ala Asp Phe Gly Asp Tyr Phe Met Pro Phe Ala Ser
            580                 585                 590
Tyr Ser Arg Thr His Arg Met Pro Asn Ile Gln Glu Met Tyr Phe Ser
        595                 600                 605
Gln Ile Gly Asp Ser Gly Val His Thr Ala Leu Lys Pro Glu Arg Ala
    610                 615                 620
Asn Thr Trp Gln Phe Gly Phe Asn Thr Tyr Lys Lys Gly Leu Leu Lys
625                 630                 635                 640
Gln Asp Asp Thr Leu Gly Leu Lys Leu Val Gly Tyr Arg Ser Arg Ile
                645                 650                 655
Asp Asn Tyr Ile His Asn Val Tyr Gly Lys Trp Trp Asp Leu Asn Gly
            660                 665                 670
Asp Ile Pro Ser Trp Val Ser Ser Thr Gly Leu Ala Tyr Thr Ile Gln
        675                 680                 685
His Arg Asn Phe Lys Asp Lys Val His Lys His Gly Phe Glu Leu Glu
    690                 695                 700
Leu Asn Tyr Asp Tyr Gly Arg Phe Phe Thr Asn Leu Ser Tyr Ala Tyr
705                 710                 715                 720
Gln Lys Ser Thr Gln Pro Thr Asn Phe Ser Asp Ala Ser Glu Ser Pro
                725                 730                 735
Asn Asn Ala Ser Lys Glu Asp Gln Leu Lys Gln Gly Tyr Gly Leu Ser
            740                 745                 750
Arg Val Ser Ala Leu Pro Arg Asp Tyr Gly Arg Leu Glu Val Gly Thr
        755                 760                 765
Arg Trp Leu Gly Asn Lys Leu Thr Leu Gly Gly Ala Met Arg Tyr Phe
    770                 775                 780
Gly Lys Ser Ile Arg Ala Thr Ala Glu Glu Arg Tyr Ile Asp Gly Thr
785                 790                 795                 800
Asn Gly Gly Asn Thr Ser Asn Phe Arg Gln Leu Gly Lys Arg Ser Ile
```

805 810 815
Lys Gln Thr Glu Thr Leu Ala Arg Gln Pro Leu Ile Phe Asp Phe Tyr
820 825 830
Ala Ala Tyr Glu Pro Lys Lys Asn Leu Ile Phe Arg Ala Glu Val Lys
835 840 845
Asn Leu Phe Asp Arg Arg Tyr Ile Asp Pro Leu Asp Ala Gly Asn Asp
850 855 860
Ala Ala Thr Gln Arg Tyr Tyr Ser Ser Phe Asp Pro Lys Asp Lys Asp
865 870 875 880
Glu Asp Val Thr Cys Asn Ala Asp Lys Thr Leu Cys Asn Gly Lys Tyr
885 890 895
Gly Gly Thr Ser Lys Ser Val Leu Thr Asn Phe Ala Arg Gly Arg Thr
900 905 910
Phe Leu Ile Thr Met Ser Tyr Lys Phe
915 920

<210> 7
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 7
gctagctagc agatcttctt tcgggaagaa 30

<210> 8
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 8
ggtcgctcga gaaacttgta gctcatcgtt atca 34

<210> 9
<211> 1000
<212> DNA

<213> Bacteria

<400> 9

tatccgataa gtttccgtac cgaacagact agattcccgc ctgcgcggga atgacgattc 60
ataagtttcc cgaaatttcca acataaccga aacctgacag taaccgtagc aactgaaccg 120
tcattcccac gaaagtggga atctagaaat gaaaagcaac aggcatttat cggaaataac 180
tgaaaccgaa cagactagat tcccgcctgc gcgggaatga cgattcataa gtttcccgaa 240
attccaacat aaccgaaacc tgacagtaac cgtagtaact gaaccgtcat tcccacgaaa 300
gtgggaatct agaaaatgaaa agcaacaggc atttatcgga aataactgga accgaacaga 360
cctagattcc cgcctgcgcg ggaatgacgg ctgcagatgc ccgacggtct ttatagcaga 420
ttaacaaaaa tcaggacaag gcggcgagcc acagacagta caaacagtac ggaaccgatt 480
cacttggtgc ttcagcacct tagagaatcg ttctctttga gctaaggcga gacaacgccg 540
tactggtttt tgttaatccg ctatattccg ccatctctaa gatttacagc gatacacggg 600
taatttaagg aatggccgaa ccgtcattcc cgccactttt cgtcattccc accacttttc 660
gtcattcccg cgcaggcggg aatctagaat ctcggacttt cagataatct ttgaatattg 720
ctgttgttct aaggtttaga ttccccttgc gcgggaatga cgattcataa gtttcccgaa 780
attccaacat aaccgaaacc tgacagtaac cgtagcaaca gagaatcgtt ctctttgagc 840
taaggcgaaa caacgctgta ccggtttttg ttaatccact ataaatatcc aattgaaatc 900
ttcagacggt atatcgaatt tacacttttt aatgtttatg ccgcctaaaa aaatgctaat 960
atatttctta attgtctgac tgtttattgt tgaggaaaat 1000

SEQUENCE INFORMATION

BASB024 Polynucleotide and Polypeptide Sequences

SEQ ID NO:1
Neisseria meningitidis **BASB024 polynucleotide sequence from strain ATCC 13090**

ATGAGATCTTCTTTCCGGTTGAAGCCGATTTGTTTTTATCTTATGGGTGTTATGCTATATCATCATAGTTATGCCGAAGA
TGCAGGGCGGCGCGGGCAGCGAGGCGCAGATACAGGTTTTGGAAGATGTGCACGTCAAGGCGAAGCGCGTACCGAAAGACA
AAAAAGTGTTTACCGATGCGCGTGCCGTATCGACCCGTCAGGATATATTCAAATCCAGCGAAAACCTCGACAACATCGTA
CGCAGCATCCCCGGTGCGTTTACACAGCAAGATAAAAGCTCGGGCATTGTGTCTTTGAATATTCGCGGCGACAGCGGGTT
CGGGCGGGTCAATACGATGGTGGACGGCATCACGCAGACCTTTTATTCGACTTCTACCGATGCGGGCAGGGCAGGCGGTT
CATCTCAATTCGGTGCATCTGTCGACAGCAATTTTATTGCCGGACTGGATGTCGTCAAAGGCAGCTTCAGCGGCTCGGCA
GGCATCAACAGCCTTGCCGGTTCGGCGAATCTGCGGACTTTAGGCGTGGATGACGTCGTTCAGGGCAATAATACCTACGG
CCTGCTGCTAAAAGGTCTGACCGGCACCAATTCAACCAAAGGTAATGCGATGGCGGCGATAGGTGCGCGCAAATGGCTGG
AAAGCGGAGCATCTGTCGGTGTGCTTTACGGGCACAGCAGGCGCACGTGGGCGCAAAATTACCGCGTGGGCGGCGGCGGG
CAGCACATCGGAAATTTTGGCGCGGAATATCTGGAACGGCGCAAACAGCGATATTTTGTACAAGAAGGCGGGTTGAAATT
CAATTCCAACAGCGGAAAATGGGAGCGGGATTTCCAAAGGCCGTACTGGAAAACCAAGTGGTATCAAAAATACAATGACC
CCCAAGAACTGCAAAAATACATCGAAGGTCATGACAAAAGCTGGCGGGAAAACCTGGCGCCGCAATACGACATCACCCCC
ATCGATCCGTCCAGCCTGAAGCAGCAGTCGGCAGGCAATCTGTTTAAATTGGAATACGACGGCGTATTCAATAAATACAC
GGCGCAATTTCGCGATTTAAACACCAAAATCGGCAGCCGCAAAATCATCAACCGCAATTATCAATTCAATTACGGTTTAT
CTTTAAACTCATATGCCAACCTCAATCTGACCGCAGCCTACAATTCGGGCAGGCAGAAATATCCGAAAGGGTCGAAGTTT
ACAGGCTGGGGGCTTTTAAAAGATTTTGAAACCTACAACAACGCGAAAATCCTCGACCTCAACAACACCGCCACCTTCCG
GCTGCCCCGCGAAACCGAGTTGCAAACCACTTTGGGCTTCAATTATTTCCACAACGAATACGGCAAAAACCGCTTTCCTG
AAGAATTGGGGCTGTTTTTCGACGGTCCGGATCAGGACAACGGGCTTTATTCCTATTTGGGGCGGTTTAAGGGCGATAAA
GGGCTGCTGCCCCAAAAATCAACCATCGTCCAACCGGCCGGCAGCCAATATTTCAACACGTTCTACTTCGATGCCGCGCT
CAAAAAAGACATTTACCGCTTAAACTACAGCACCAATACCGTCGGCTACCGTTTCGGCGGCGAATATACGGGCTATTACG
GCTCGGATGACGAATTTAAGCGGGCATTCGGAGAAAACTCGCCGACATACAAGAAACATTGCAACCAGAGCTGCGGAATT
TATGAACCCGTATTGAAAAAATACGGCAAAAAGCGCGCCAACAACCATTCGGTCAGCATTAGTGCGGACTTCGGCGATTA
TTTCATGCCGTTCGCCAGCTATTCGCGCACACACCGTATGCCCAACATCCAAGAAATGTATTTTTCCCAAATCGGCGACT
CCGGCGTTCACACCGCCTTAAAACCAGAGCGCGCAAACACTTGGCAATTTGGCTTCAATACCTATAAAAAAGGATTGTTA
AAACAAGATGATACATTAGGATTAAAACTGGTCGGCTACCGCAGCCGCATCGACAACTACATCCACAACGTTTACGGGAA
ATGGTGGGATTTGAACGGGAATATTCCGAGCTGGGTCAGCAGCACCGGGCTTGCCTACACCATCCAACACCGCAATTTCA
AAGACAAAGTACACAAACACGGTTTTGAGTTGGAGCTGAATTACGATTATGGGCGTTTTTTCACCAACCTTTCTTACGCC
TATCAAAAAAGCACGCAACCGACCAACTTCAGCGATGCGAGCGAATCGCCCAACAATGCGTCCAAAGAAGACCAACTCAA
ACAAGGTTATGGGTTGAGCAGGGTTTCCGCCCTGCCGCGAGATTACGGACGTTTGGAAGTCGGTACGCGCTGGTTGGGCA
ACAAACTGACTTTGGGCGGCGCGATGCGCTATTTCGGCAAGAGCATCCGCGCGACGGCTGAAGAACGCTATATCGACGGC
ACCAACGGGGGAAATACCAGCAATGTCCGGCAACTGGGCAAGCGTTCCATCAAACAAACCGAAACCCTTGCCGCCAGCC
TTTGATTTTTGATTTTTACGCCGCTTACGAGCCGAAGAAAAACCTTATTTTCCGCGCCGAAGTCAAAAATCTGTTCGACA

GGCGTTATATCGATCCGCTCGATGCGGGCAATGATGCGGCAACGCAGCGTTATTACAGTTCGTTCGACCCGAAAGACAAG
GACGAAGAAGTAACGTGTAATGCTGATAAAACGTTGTGCAACGGCAAATACGGCGGCACAAGCAAAAGCGTATTGACCAA
TTTTGCACGCGGACGCACCTTTTTGATAACGATGAGCTACAAGTTTTAA

SEQ ID NO:2

Neisseria meningitidis **BASB024 polypeptide sequence deduced from the polynucleotide of SeQ ID NO:1**

MRSSFRLKPICFYLMGVMLYHHSYAEDAGRAGSEAQIQVLEDVHVKAKRVPKDKKVFTDARAVSTRQDIFKSSENLDNIV
RSIPGAFTQQDKSSGIVSLNIRGDSGFGRVNTMVDGITQTFYSTSTDAGRAGGSSQFGASVDSNFIAGLDVVKGSFSGSA
GINSLAGSANLRTLGVDDVVQGNNTYGLLLKGLTGTNSTKGNAMAAIGARKWLESGASVGVLYGHSRRTWAQNYRVGGGG
QHIGNFGAEYLERRKQRYFVQEGGLKFNSNSGKWERDFQRPYWKTKWYQKYNDPQELQKYIEGHDKSWRENLAPQYDITP
IDPSSLKQQSAGNLFKLEYDGVFNKYTAQFRDLNTKIGSRKIINRNYQFNYGLSLNSYANLNLTAAYNSGRQKYPKGSKF
TGWGLLKDFETYNNAKILDLNNTATFRLPRETELQTTLGFNYFHNEYGKNRFPEELGLFFDGPDQDNGLYSYLGRFKGDK
GLLPQKSTIVQPAGSQYFNTFYFDAALKKDIYRLNYSTNTVGYRFGGEYTGYYGSDDEFKRAFGENSPTYKKHCNQSCGI
YEPVLKKYGKKRANNHSVSISADFGDYFMPFASYSRTHRMPNIQEMYFSQIGDSGVHTALKPERANTWQFGFNTYKKGLL
KQDDTLGLKLVGYRSRIDNYIHNVYGKWWDLNGNIPSWVSSTGLAYTIQHRNFKDKVHKHGFELELNYDYGRFFTNLSYA
YQKSTQPTNFSDASESPNNASKEDQLKQGYGLSRVSALPRDYGRLEVGTRWLGNKLTLGGAMRYFGKSIRATAEERYIDG
TNGGNTSNVRQLGKRSIKQTETLARQPLIFDFYAAYEPKKNLIFRAEVKNLFDRRYIDPLDAGNDAATQRYYSSFDPKDK
DEEVTCNADKTLCNGKYGGTSKSVLTNFARGRTFLITMSYKF

SEQ ID NO:3

Neisseria meningitidis **BASB024 polynucleotide sequence from strain ATCC 13090**

ATGAGATCTTCTTTCCGGTTGAAGCCGATTTGTTTTTATCTTATGGGTGTTATGCTATATCATCATAGTTATGCCGAAGA
TGCAGGGCGCGCGGGCAGCGAGGCGCAGATACAGGTTTGGAAGATGTGCACGTCAAGGCGAAGCGCGTACCGAAAGACA
AAAAAGTGTTTACCGATGCGCGTGCCGTATCGACCCGTCAGGATATATTCAAATCCAGCGAAAACCTCGACAACATCGTA
CGCAGCATCCCCGGTGCGTTTACACAGCAAGATAAAAGCTCGGGCATTGTGTCTTTGAATATTCGCGGCGACAGCGGGTT
CGGGCGGGTCAATACGATGGTGGACGGCATCACGCAGACCTTTTATTCGACTTCTACCGATGCGGGCAGGGCAGGCGGTT
CATCTCAATTCGGTGCATCTGTCGACAGCAATTTTATTGCCGGACTGGATGTCGTCAAAGGCAGCTTCAGCGGCTCGGCA
GGCATCAACAGCCTTGCCGGTTCGGCGAATCTGCGGACTTTAGGCGTGGATGACGTCGTTCAGGGCAATAATACCTACGG
CCTGCTGCTAAAAGGTCTGACCGGCACCAATTCAACCAAAGGTAATGCGATGGCGGCGATAGGTGCGCGCAAATGGCTGG
AAAGCGGAGCATCTGTCGGTGTGCTTTACGGGCACAGCAGGCGCAgcgtGGCGCAAAATTACCGCGTGGGCGGCGGCGGG
CAGCACATCGGAAATTTTGGCGCGGAATATCTGGAACGGCGCAAgCAGCGATATTTTGTACAAGAAGGCGGGTTGAAATT
CAATTCCAACAGCGGAAAATGGGAGCGGGATTTCCAAAGGCCGTACTGGAAAACCAAGTGGTATCAAAAATACAATGACC
CCCAAGAACTGCAAAAATACATCGAAGGTCATGACAAAAGCTGGCGGGAAAACCTGGCGCCGCAATACGACATCACCCCC
ATCGATCCGTCCAGCCTGAAGCAGCAGTCGGCAGGCAATCTGTTTAAATTGGAATACGACGGCGTATTCAATAAATACAC
GGCGCAATTTCGCGATTTAAACACCAAAATCGGCAGCCGCAAAATCATCAACCGCAATTATCAATTCAATTACGGTTTAT
CTTTAAACTCATATGCCAACCTCAATCTGACCGCAGCCTACAATTCGGGCAGGCAGAAATATCCGAAAGGGTCGAAGTTT

ACAGGCTGGGGGCTTTTAAAAGATTTTGAAACCTACAACAACGCGAAAATCCTCGACCTCAACAACACCGCCACCTTCCG
GCTGCCCCGCGAAACCGAGTTGCAAACCACTTTGGGCTTCAATTATTTCCACAACGAATACGGCAAAAACCGCTTTCCTG
AAGAATTGGGGCTGTTTTTCGACGGTCCGGATCAGGACAACGGGCTTTATTCCTATTTGGGGCGGTTTAAGGGCGATAAA
GGGCTGCTGCCCCAAAAATCAACCATCGTCCAACCGGCCGGCAGCCAATATTTCAACACGTTCTACTTCGATGCCGCGCT
CAAAAAAGACATTTACCGCTTAAACTACAGCACCAATACCGTCGGCTACCGTTTCGGCGGCGAATATACGGGCTATTACG
GCTCGGATGACGAATTTAAGCGGGCATTCGGAGAAAACTCGCCGACATACAAGAAACATTGCAACCAGAGCTGCGGAATT
TATGAACCCGTATTGAAAAAATACGGCAAAAAGCGCGCCAACAACCATTCGGTCAGCATTAGTGCGGACTTCGGCGATTA
TTTCATGCCGTTCGCCAGCTATTCGCGCACACACCGTATGCCCAACATCCAAGAAATGTATTTTTCCCAAATCGGCGACT
CCGGCGTTCACACCGCCTTAAAACCAGAGCGCGCAAACACTTGGCAATTTGGCTTCAATACCTATAAAAAAGGATTGTTA
AAACAAGATGATACATTAGGATTAAAACTGGTCGGCTACCGCAGCCGCATCGACAACTACATCCACAACGTTTACGGGAA
ATGGTGGGATTTGAACGGGAATATTCCGAGCTGGGTCAGCAGCACCGGGCTTGCCTACACCATCCAACACCGCAATTTCA
AAGACAAAGTACACAAACACGGTTTTGAGTTGGAGCTGAATTACGATTATGGGCGTTTTTTCACCAACCTTTCTTACGCC
TATCAAAAAAGCACGCAACCGACCAACTTCAGCGATGCGAGCGAATCGCCCAACAATGCGTCCAAAGAAGACCAACTCAA
ACAAGGTTATGGGTTGAGCAGGGTTTCCGCCCTGCCGCGAGATTACGGACGTTTGGAAGTCGGTACGCGCTGGTTGGGCA
ACAAACTGACTTTGGGCGGCGCGATGCGCTATTTCGGCAAGAGCATCCGCGCGACGGCTGAAGAACGCTATATCGACGGC
ACCAACGGGGGAAATACCAGCAATGTCCGGCAACTGGGCAAGCGTTCCATCAAACAAACCGAAACCCTTGCCCGCCAGCC
TTTGATTTTTGATTTTTACGCCGCTTACGAGCCGAAGAAAAACCTTATTTTCCGCGCCGAAGTCAAAAATCTGTTCGACA
GGCGTTATATCGATCCGCTCGATGCGGGCAATGATGCGGCAACGCAGCGTTATTACAGTTCGTTCGACCCGAAAGACAAG
GACGAAGAAGTAACGTGTAATGCTGATAAAACGTTGTGCAACGGCAAATACGGCGGCACAAGCAAAAGCGTATTGACCAA
TTTTGCACGCGGACGCACCTTTTTGATAACGATGAGCTACAAGTTTTAA

SEQ ID NO:4
***Neisseria meningitidis* BASB024 polypeptide sequence deduced from the polynucleotide of SeQ ID NO:3**

MRSSFRLKPICFYLMGVMLYHHSYAEDAGRAGSEAQIQVLEDVHVKAKRVPKDKKVFTDARAVSTRQDIFKSSENLDNIV
RSIPGAFTQQDKSSGIVSLNIRGDSGFGRVNTMVDGITQTFYSTSTDAGRAGGSSQFGASVDSNFIAGLDVVKGSFSGSA
GINSLAGSANLRTLGVDDVVQGNNTYGLLLKGLTGTNSTKGNAMAAIGARKWLESGASVGVLYGHSRRSVAQNYRVGGGG
QHIGNFGAEYLERRKQRYFVQEGGLKFNSNSGKWERDFQRPYWKTKWYQKYNDPQELQKYIEGHDKSWRENLAPQYDITP
IDPSSLKQQSAGNLFKLEYDGVFNKYTAQFRDLNTKIGSRKIINRNYQFNYGLSLNSYANLNLTAAYNSGRQKYPKGSKF
TGWGLLKDFETYNNAKILDLNNTATFRLPRETELQTTLGFNYFHNEYGKNRFPEELGLFFDGPDQDNGLYSYLGRFKGDK
GLLPQKSTIVQPAGSQYFNTFYFDAALKKDIYRLNYSTNTVGYRFGGEYTGYYGSDDEFKRAFGENSPTYKKHCNQSCGI
YEPVLKKYGKKRANNHSVSISADFGDYFMPFASYSRTHRMPNIQEMYFSQIGDSGVHTALKPERANTWQFGFNTYKKGLL
KQDDTLGLKLVGYRSRIDNYIHNVYGKWWDLNGNIPSWVSSTGLAYTIQHRNFKDKVHKHGFELELNYDYGRFFTNLSYA
YQKSTQPTNFSDASESPNNASKEDQLKQGYGLSRVSALPRDYGRLEVGTRWLGNKLTLGGAMRYFGKSIRATAEERYIDG
TNGGNTSNVRQLGKRSIKQTETLARQPLIFDFYAAYEPKKNLIFRAEVKNLFDRRYIDPLDAGNDAATQRYYSSFDPKDK
DEEVTCNADKTLCNGKYGGTSKSVLTNFARGRTFLITMSYKF

SEQ ID NO:5
***Neisseria meningitidis* BASB024 polynucleotide sequence from strain H44/76**

ATGAGATCTTCTTTCCGGTTGAAGCCGATTTGTTTTTAcCTTATGGGTGTTAcGCTATATCATcATAGTTATGCCGAAGA
TGCAGGGCGCGCGGGCAGCGAGGCGCAGATACAGGTTTTGGAAGATGTGCACGTCAAGGCGAAGCGCGTACCGAAAGACA
AAAAAGTGTTTACCGATGCGCGTGCCGTATCGACCCGTCAGGATATATTCAAATCCAGCGAAAACCTCGACAACATCGTA
CGCAGCATCCCCGGTGCGTTTACACAGCAAGATAAAAGCTCGGGCATTGTGTCTTTGAATATTCGCGGCGACAGCGGGTT
CGGGCGGGTCAATACGATGGTGGACGGCATCACGCAGACCTTTATTCGACTTCTACCGATGCGGGCAGGGCAGGCGGTT
CATCTCAATTCGGTGCATCTGTCGACAGCAATTTTATTGCCGGACTGGATGTCGTCAAAGGCAGCTTCAGCGGCTCGGCA
GGCATCAACAGCCTTGCCGGTTCGGCGAATCTGCGGACTTTAGGCGTGGATGACGTCGTTCAGGGCAATAATACCTACGG
CCTGCTGCTAAAAGGTCTGACCGGCACCAATTCAACCAAAGGTAATGCGATGGCGGCGATAGGTGCGCGCAAATGGCTGG
AAAGCGGAGCATCTGTCGGTGTGCTTTACGGGCACAGCAGGCGCAGCGTGGCGCAAAATTACCGCGTGGGCGGCGGCGGG
CAGCACATCGGAAATTTTGGCGCGGAATATTTGGAACGGCGCAAGCAGCGATATTTTGTACAAGAGGGTGCTTTGAAATT
CAATTCCGACAGCGGAAAATGGGAGCGGGATTTACAAAGGCAACAGTGGAAATACAAGCCGTATAAAAATTACAACAACC
AAGAACTACAAAAATACATCGAAGGTCATGACAAAAGCTGGCGGGAAAACCTGGCGCCGCAATACGACATCACCCCCATC
GATCCGTCCAGCCTGAAGCAgcagtcggcAGGCAATCTGTTTAAATTGGAATACGACGGCGTATTCAATAAATACACGGC
GCAATTTCGCGATTTAAACACCAAAATCGGCAGCCGCAAAATCATCAACCGCAATTATCAgTTCAATTACGGTTTgTCTT
TgAACcCgTATaCCAACCTCAATCTGACCGCAGCCTACAATTCGGGCAGGCAGAAATATCCGAAAGGGTCGAAGTTTACA
GGCTGGGGGCTTTTAAAGGATTTTGAAACCTACAACAACGCGAAAATCCTCGACCTCAACAACACCGCCACCTTCCGGCT
GCCCCGCGAAACCGAGTTGCAAACCACTTTGGGCTTCAATTATTTCCACAACGAATACGGCAAAAACCGCTTTCCTGAAG
AATTGGGGCTGTTTTTCGACGGTCCTGATCAGGACAACGGGCTTTATTCCTATTTGGGGCGGTTTAAGGGCGATAAAGGG
CTGCTGCCCCAAAAATCAACCATTGTCCAACCGGCCGGCAGCCAATATTTCAACACGTTCTACTTCGATGCCGCGCTCAA
AAAAGACATTTACCGCTTAAACTACAGCACCAATACCGTCGGCTACCGTTTCGGCGGCGAATATACGGGCTATTACGGCT
CGGATGACGAATTTAAGCGGGCATTCGGAGAAAACTCGCCGACATACAAGAAACATTGCAACCGGAGCTGCGGGATTTAT
GAACCCGTATTGAAAAAATACGGCAAAAAGCGCGCCAACAACCATTCGGTCAGCATTAGTGCGGACTTCGGCGATTATTT
CATGCCGTTCGCCAGCTATTCGCGCACACACCGTATGCCCAACATCCAAGAAATGTATTTTTCCCAAATCGGCGACTCCG
GCGTTCACACCGCCTTAAAACCAGAGCGCGCAAACACTTGGCAATTTGGCTTCAATACCTATAAAAAAGGATTGTTAAAA
CAAGATGATACATTAGGATTAAAACTGGTCGGCTACCGCAGCCGCATCGACAACTACATCCACAACGTTTACGGGAAATG
GTGGGATTTGAACGGGgATATTCCGAGCTGGGTCAGCAGCACCGGGCTTGCCTACACCATCCAACATCGCAATTTCAAAG
ACAAAGTGCACAAACACGGTTTTGAGTTGGAGCTGAATTACGATTATGGGCGTTTTTTCACCAACCTTTCTTACGCCTAT
CAAAAAAGCACGCAACCGACCAACTTCAGCGATGCGAGCGAATCGCCCAACAATGCGTCCAAAGAAGACCAACTCAAACA
AGGTTATGGGTTGAGCAGGGTTTCCGCCCTGCCGCGAGATTACGGACGTTTGGAAGTCGGTACGCGCTGGTTGGGCAACA
AACTGACTTTGGGCGGCGGCGATGCGCTATTTCGGCAAGAGCATCCGCGCGACGGCTGAAGAACGCTATATCGACGGCACC
AACGGGGGAAATACCAGCAATTTCCGGCAACTGGGCAAGCGTTCCATCAAACAAACCGAAACTCTTGCCCGCCAGCCTTT
GATTTTTGATTTTTACGCCGCTTACGAGCCGAAGAAAAACCTTATTTTCCGCGCCGAAGTCAAAAATCTGTTCGACAGGC
GTTATATCGATCCGCTCGATGCGGGCAATGATGCGGCAACGCAGCGTTATTACAGCTCGTTCGACCCGAAAGACAAGGAC
GAAGACGTAACGTGTAATGCTGATAAAACGTTGTGCAACGGCAAATACGGCGGCACAAGCAAAAGCGTATTGACCAATTT
TGCACGCGGACGCACCTTTTTGATAACGATGAGCTACAAGTTTTAA

SEQ ID NO:6
***Neisseria meningitidis* BASB024 polypeptide sequence deduced from the polynucleotide of SeQ ID NO:5**

MRSSFRLKPICFYLMGVTLYHYSYAEDAGRAGSEAQIQVLEDVHVKAKRVPKDKKVFTDARAVSTRQDIFKSSENLDNIV
RSIPGAFTQQDKSSGIVSLNIRGDSGFGRVNTMVDGITQTFYSTSTDAGRAGGSSQFGASVDSNFIAGLDVVKGSFSGSA
GINSLAGSANLRTLGVDDVVQGNNTYGLLLKGLTGTNSTKGNAMAAIGARKWLESGASVGVLYGHSRRSVAQNYRVGGGG
QHIGNFGAEYLERRKQRYFVQEGALKFNSDSGKWERDLQRQQWKYKPYKNYNNQELQKYIEGHDKSWRENLAPQYDITPI
DPSSLKQQSAGNLFKLEYDGVFNKYTAQFRDLNTKIGSRKIINRNYQFNYGLSLNPYTNLNLTAAYNSGRQKYPKGSKFT
GWGLLKDFETYNNAKILDLNNTATFRLPRETELQTTLGFNYFHNEYGKNRFPEELGLFFDGPDQDNGLYSYLGRFKGDKG
LLPQKSTIVQPAGSQYFNTFYFDAALKKDIYRLNYSTNTVGYRFGGEYTGYYGSDDEFKRAFGENSPTYKKHCNRSCGIY
EPVLKKYGKKRANNHSVSISADFGDYFMPFASYSRTHRMPNIQEMYFSQIGDSGVHTALKPERANTWQFGFNTYKKGLLK
QDDTLGLKLVGYRSRIDNYIHNVYGKWWDLNGDIPSWVSSTGLAYTIQHRNFKDKVHKHGFELELNYDYGRFFTNLSYAY
QKSTQPTNFSDASESPNNASKEDQLKQGYGLSRVSALPRDYGRLEVGTRWLGNKLTLGGAMRYFGKSIRATAEERYIDGT
NGGNTSNFRQLGKRSIKQTETLARQPLIFDFYAAYEPKKNLIFRAEVKNLFDRRYIDPLDAGNDAATQRYYSSFDPKDKD
EDVTCNADKTLCNGKYGGTSKSVLTNFARGRTFLITMSYKF

SEQ ID NO:7
GCT AGC TAG CAG ATC TTC TTT CGG GAA GAA

SEQ ID NO:8
GGT CGC TCG AGA AAC TTG TAG CTC ATC GTT ATC A

SEQ ID NO:9
***Neisseria meningitidis* polynucleotide sequence up-stream the BASB024 gene sequence, in strain ATCC 13090**

TATCCGATAAGTTTCCGTACCGAACAGACTAGATTCCCGCCTGCGCGGGAATGACGATTCATAAGTTTCCCGAAATTCCA
ACATAACCGAAACCTGACAGTAACCGTAGCAACTGAACCGTCATTCCCACGAAAGTGGGAATCTAGAAATGAAAAGCAAC
AGGCATTTATCGGAAATAACTGAAACCGAACAGACTAGATTCCCGCCTGCGCGGGAATGACGATTCATAAGTTTCCCGAA
ATTCCAACATAACCGAAACCTGACAGTAACCGTAGTAACTGAACCGTCATTCCCACGAAAGTGGGAATCTAGAAATGAAA
AGCAACAGGCATTTATCGGAAATAACTGGAACCGAACAGACCTAGATTCCCGCCTGCGCGGGAATGACGGCTGCAGATGC
CCGACGGTCTTTATAGCAGATTAACAAAAATCAGGACAAGGCGGCGAGCCACAGACAGTACAAACAGTACGGAACCGATT
CACTTGGTGCTTCAGCACCTTAGAGAATCGTTCTCTTTGAGCTAAGGCGAGACAACGCCGTACTGGTTTTTGTTAATCCG
CTATATTCCGCCATCTCTAAGATTTACAGCGATACACGGGTAATTTAAGGAATGGCCGAACCGTCATTCCCGCCACTTTT
CGTCATTCCCACCACTTTTCGTCATTCCCGCGCAGGCGGGAATCTAGAATCTCGGACTTTCAGATAATCTTTGAATATTG
CTGTTGTTCTAAGGTTTAGATTCCCCTTGCGCGGGAATGACGATTCATAAGTTTCCCGAAATTCCAACATAACCGAAACC
TGACAGTAACCGTAGCAACAGAGAATCGTTCTCTTTGAGCTAAGGCGAAACAACGCTGTACCGGTTTTTGTTAATCCACT
ATAAATATCCAATTGAAATCTTCAGACGGTATATCGAATTTACACTTTTTAATGTTTATGCCGCCTAAAAAAATGCTAAT
ATATTTCTTAATTGTCTGACTGTTTATTGTTGAGGAAAAT